PUSHING THE CELL ENVELOPE: SIMULATIONS AT THE GRAM-NEGATIVE CELLULAR INTERFACE

A Dissertation
Presented to
The Academic Faculty

By

Curtis A. Balusek

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Physics

Georgia Institute of Technology

May 2019

PUSHING THE CELL ENVELOPE: SIMULATIONS AT THE GRAM-NEGATIVE CELLULAR INTERFACE

Approved by:

Dr. James C. Gumbart, Advisor
School of Physics
Georgia Institute of Technology

Dr. Jennifer Curtis
School of Physics
Georgia Institute of Technology

Dr. Peter Yunker
School of Physics
Georgia Institute of Technology

Dr. Jesse McDaniels
School of Chemistry
Georgia Institute of Technology

Dr. Loren Williams
School of Chemistry
Georgia Institute of Technology

Date Approved: December 12, 2018

"There are two reasons, and two reasons alone for mathematics. The first is so you don't get f#!$ed over in a business transaction. The second, and more important reason, is to help mankind describe the natural world, and that is called physics.

-Mr. Raven, 12th Grade Physics

First, to my friends, for being there when I needed to relax, have a beer, or talk it over. Second to my parents, for their steadfast support in my educational endeavors. And lastly, but most crucially, to my love Linda; I would not have made it this far if it wasn't for all of your love and support.

ACKNOWLEDGEMENTS

There comes a time when you stop to reflect on all the people who have contributed to who you have become and how these people have shaped your life. There have been many along this road towards completing my degree and it starts back in Undergraduate.

I recall my first steps onto Sam Houston State, and decided there was no way I was going to take a Floral Design course to get an Animal Science degree. I walked over to the math building and met with the first person who would steer my course to where I am today. Prof. Jacquelyn Jensen, was the undergraduate mathematics advisor and she told me that in no few terms that the likelihood I would get a PhD was slim if not impossible. Challenge accepted. Her initial push guided me towards completing a Physics and Math dual degree shaped by Prof. John Alford, who I worked extensively with on creating applied math models of hydrilla and snakes. Working with Prof. Barry Friedman in the school of physics shaped my computational prowess by working with COMSOL and even some FORTRAN77.

There was also my involvement with the Honors college and the friends who changed how I view and envision the world and the people within in. Working with Prof. Gene Young, the then director of the Honors Program as well as Dr. Maria Holmes to shape what would become the Honors College. The numerous friends within the college whom I am still friends with to this day: Ashley, Joe, Kodi, Chris, Beth, Brittany, many of whom have gone on to complete their MS or PhD: Forensic Sci., Forensic Sci., Library Sci., Forensic Chemistry, Entomology, and Toxicology respectively. Each of the people from my undergraduate helped guide my own destiny towards Georgia Tech and becoming a PhD in physics.

Once I arrived at Tech, it was a change of pace and of life. The tumultuous, rigorous coursework of the first year would have been impossible to accomplish alone and if it weren't for Gable and Tom, it would have been that much worse. We were a large class

of misfits, but we each had our personalities. Then into my research group. I was the first student to work with Dr. JC Gumbart, and through my coercion, managed to find a few colleagues to join, Anthony and later Karl. JC drove a hard ship, there were good time and bad; rocky roads, that seemed to stretch on forever, but we all became better scientists for the amount of hard work and difficulty we had, willingly or not. I thank Anthony for transforming the energy landscape within the lab, literally, with his work on peptide folding. I also thank Karl for his efforts in describing the OM environment that much of the work contained within was built around. I would also like to thank Sunny for her work on membranes as well as her contagious dedication to getting work done.

There were many others during my time here at Tech, Coach Estrada and his family who took in a stray and helped feed, shelter, and encourage through each of the twists in the metaphorical road of PhD work. To my extended archery family for helping support and encourage me throughout this entire process by creating an environment of curiosity and assistance.

I also want to acknowledge my scientific friends for each of their conversations, which in their own way helped shape a lot of my work as well as helped me find my way through dark times. To my friends in the school of Chemistry, Biology, Quantitative Bio-sciences, Materials Science, and more, here at Tech and abroad. Each scientist stands on the shoulders of giants, but you don't realize how often those giants stand on your shoulders as well.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SYMBOLS AND ABBREVIATIONS

AA	All-atom
CG	Coarse Grain
CHARMM	Chemistry at Harvard Macromolecular Mechanics software & force-field
colvar	collective-variable
cryo-EM	cryo-Electron Microscopy
DP+	two palmitoyl (16:0) aliphatic tails
GAL	Galactose
GLC	Glucose
HEP	Heptose
HMR	Hydrogen Mass Repartitioning
IM	Inner Membrane
KDO	3-Deoxy-D-manno-oct-2-ulosonic acid or keto-deoxyoctulosonate
KCl	Potassium Chloride
LIPA	Lipid-A
LJ	Lennard-Jones
LPS	Lipopolysaccharide
MD	Molecular Dynamics
NAMD	Software for nanoscale Molecular Dynamics
NMR	Nuclear Magnetic Resonance
NPT	Constant Number, Pressure, and Temperature
OM	Outer Membrane
PDB	Protein Data Bank
PME	Particle Mesh Ewald Summation

+PC	+-phosphatidylcholine; phospholipid headgroup
+PG	+-phosphatidylglycerol; phospholipid headgroup
+PE	+-phosphatidylethanolamine; phospholipid headgroup
PO+	Palmitoyl-Oleoyl (16:0,18:1) aliphatic tails
RMSD	Root Mean Square Displacement
RMSF	Root Mean Square Fluctuation
S_{CD}	Deuterium Order Parameter
SMD	Steered Molecular Dynamics
VMD	Visual Molecular Dynamics
2-12	2 fs timestep & 12Å cutoff
4-12	4 fs timestep & 12Å cutoff
4-9	4 fs timestep & 9Å cutoff

SUMMARY

The cell envelope of Gram-negative bacteria contrasts with it's positive counterpart by containing two cellular membranes separated by a cell wall. This makes Gram-negative bacteria particularly interesting especially when considering that there are no prototypical energy resources, e.g. ATP or ion gradient, at the outer-membrane (OM) such that nutrient uptake, protein folding, lipid insertion, etc. must occur in the absence cellular energy resource.

Innumerably many experimental and computational studies have been performed on cytoplasmic membrane proteins in order to determine function. Furthermore, the past decade has seen the increase in simulating OM proteins in native, asymmetric membranes to elucidate function in the absence of cellular energy. However, the function of occupants separating these two membranes has not be studied in great detail, nor have cell-envelope spanning complexes, such as the AcrAB-TolC, multi-drug efflux pump, been studied in their natural environment. Therefore, this thesis is presented in three parts, towards the construction of a Gram-negative cell-envelope model, which can be used in computational simulations to begin looking at envelope-spanning protein complexes as well as protein-protein interactions within the periplasmic space.

The first study presents a comparison of a common, experimental, symmetric membrane to a native, asymmetric membrane and the role each has on the behavior of the OM, $\beta-$barrel protein BtuB. The results from simulation suggest that there are important membrane-protein interactions in the native OM that contribute to the stability of the extracellular loops of BtuB. These loops were shown to be stabilized by bound calcium ions in previous simulation [1]. The importance of calcium in BtuB function was shown experimentally to increase substrate-binding efficiency 1000-fold [2]. While stabilization of key extracellular loops is observed due to the membrane, a distinct structural difference exist between the calcium-bound and apo (no calcium) states in the native OM. However,

the simulations also suggest that the membrane composition creates an allosteric effect that reduces the force required to extract luminal domain from within the $\beta-$barrel.

The second study presented utilizes a repartitioning of atomic masses in order to accelerate MD simulations with lipids. Due to their amphiphilic nature, lipid molecules and their associated interactions are primarily governed by the hydrophobic effect, which is best described by London-dispersion and hard-sphere repulsion. These two attributes, commonly referred to at the van Der Waals interactions, are governed in MD by the Lennard-Jones (LJ) potential. Within this study, we use hydrogen mass repartitioning (HMR), which transfers mass from a heavy parent-atom to each of the bonded hydrogens. This process results is a decrease in oscillation frequency of hydrogen associated angles and dihedrals, such that a longer time step can be utilized while performing all-atom simulations with lipids. Furthermore, HMR is applied such that the molecule mass is conserved. For our comparison, we present results from multiple all-atom lipid simulations that compare the effect of HMR, increasing integration time-step, and reduced LJ distance cutoffs to conventional MD. We find that HMR has little effect on the structural and kinetic properties of membrane systems, an increased time step slightly effects membrane fluidity and diffusion, and lastly, a shorter cutoff drastically differs across each of the measurements used.

Lastly, the two aforementioned studies prompted by the ongoing development of cell-envelope spanning complexes led to a discussion of developing a method to construct a cell envelope model resembling the periplasmic environment. In the last chapter, we present a method from requisite genomic and proteomic information to create an abundance map of cell-envelope associated proteins. Furthermore, with the use of cross-linking and mass-spectrometry experiments, specific protein-protein interactions can be considered when constructing an all-atom model.

CHAPTER 1

INTRODUCTION

1.1 Background

Bacteria are a major concern for human health across the globe due to the rapid increase in the prevalence of antibiotic resistance, e.g. methicillin-resistant *Staphlococcus aureus* or MRSA and *N. gonorrhoeae* [3]. One particular reason for this increase in resistance is that receptor mutations have developed along the extracellular surface of the bacterial cell envelope that current antibiotic treatments can no longer target. Therefore, it is of utmost interest to understand the pathways that bacteria utilize for nutrient uptake in order to develop the next generation of antibiotics.

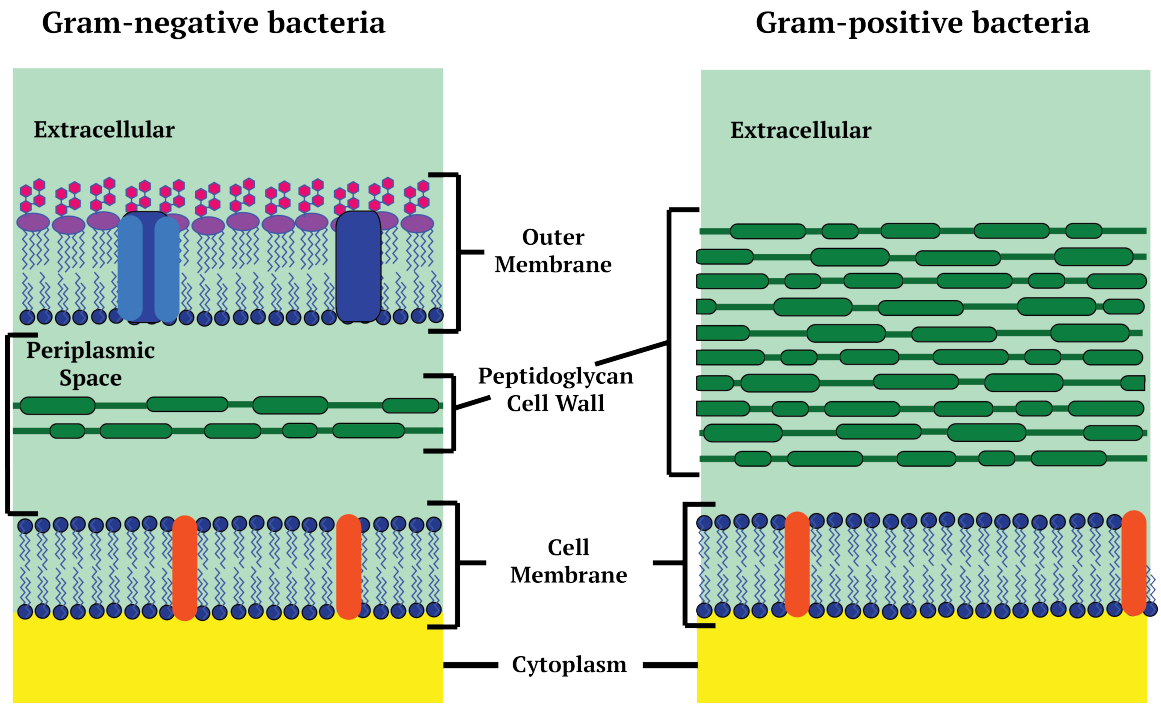


Figure 1.1: A schematic representation of Gram-negative (left) and Gram-positive (right) cell envelopes. Contrary to the positive counterpart, Gram-negative bacteria have an outer membrane surrounding the cell wall and inner membrane.

There are two primary types of bacteria, distinguished by the Gram staining process developed in the late 19^{th} century, Gram-positive and -negative, as shown in Fig. 1.1. The latter are differentiated from their positive counterparts by the presence of a second, outer membrane (OM) that surrounds the otherwise similar cell wall and inner membrane (IM). The toxicity of Gram-negative bacteria can be attributed to endotoxins, or lipopolysaccharides (LPS), that constitute part of the OM environment [4]. Unlike the prototypical lipid bilayer, the OM is an asymmetric bilayer composed of an inner leaflet of phospholipids and an outer leaflet of LPS [5, 6], shown in Fig. 1.2A. In contrast to the two aliphatic tails characteristic of phospholipids, LPS has a varying number of lipid tails, typically four to six depending on the species, as well as a large, charged oligosaccharide head group and long O-antigen chains [7, 6, 8, 9, 10], shown in Fig. 1.2B. Divalent cations promote ionic bridging between phosphate groups on the LPS core-oligosaccharides that has been predicted to reduce the lateral mobility of LPS [11], and induce a barrier to both hydrophobic and hydrophilic molecules [12, 13, 6]. Due to the reduced permeability of the LPS-containing bilayer, Gram-negative bacteria have an assortment of trans-membrane proteins that facilitate the passive diffusion or transport of molecules across the OM and into the periplasmic space between inner and outer membranes [14]. Small molecules and ions (< 600 Da) can passively diffuse through a variety of porins, or open channels in the OM; whereas large molecules and nutrients are required to pass through transporters. Characteristic of most OM proteins is a β-barrel structure spanning the membrane as well as long and charged EC loops, see Fig. 1.2C. The orientation of these EC loops often play a significant role in nutrient uptake across the OM [15, 16, 17, 18, 19]. Because ions and nutrients are required to travel through OM proteins, understanding the function and transport mechanisms of these proteins can lead to the development of the next generation of antibiotics.

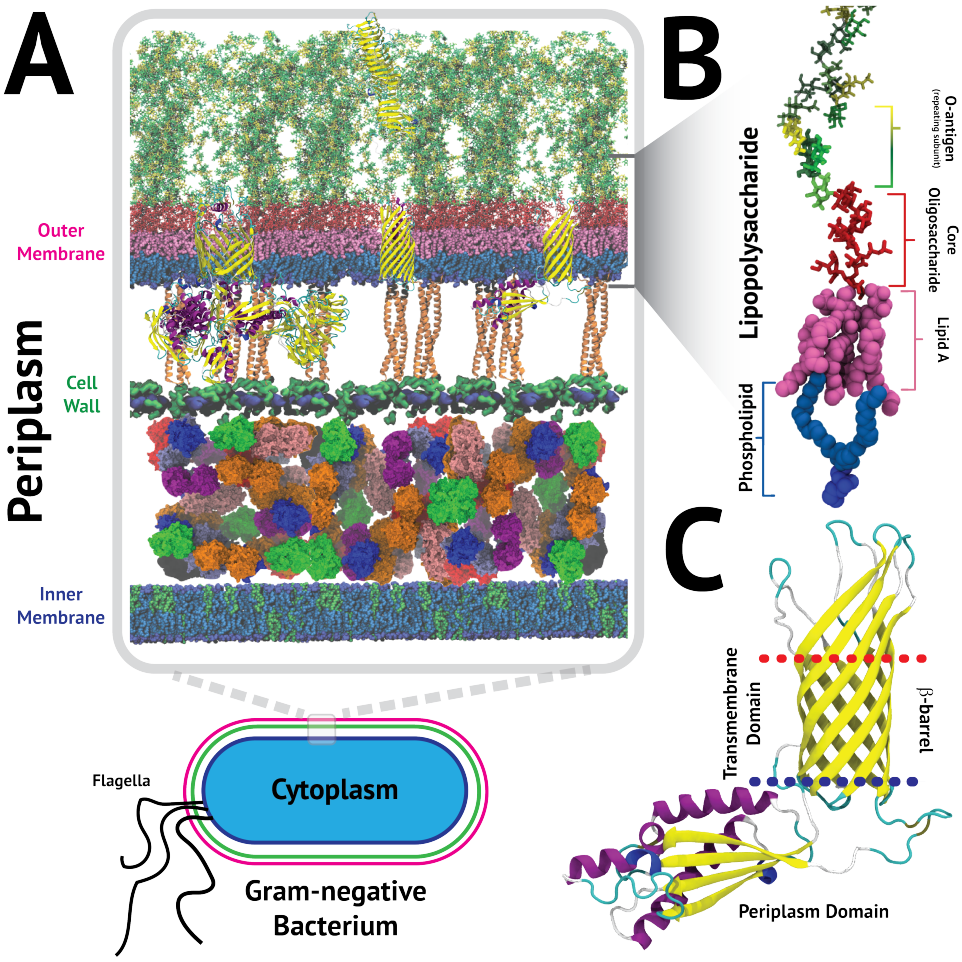


Figure 1.2: Bottom Left, sketch of a prototypical Gram-negative bacterium with the cytoplasm shown in light blue and cell envelope colors corresponding to the the model shown in A) Gram-negative cell envelope model with inner membrane (blue and green spheres), cell wall (blue and green volumes), and outer membrane (blue and magenta spheres). B) Comparison of Lipopolysaccharide (top, spheres and sticks) with phospholipid (bottom, blue spheres). The core lipid-A molecule is shown in Magenta Spheres, core oligosaccharide in red sticks, and O-antigen repeat in yellow/green sticks. C) Most abundant outer transmembrane protein in *E. coli*, OmpA; $\beta-$barrel shape (yellow) spans the OM. Red and blue dotted lines are amphiphilic interfaces with LipidA (red) and phospholipids (blue). Periplasm domain is thought to interact with the cell wall.

1.2 TonB-dependent Transporters

The open nature of porins negates the utility of cytoplasmic energy sources in the periplasm, e.g. ATP, GTP, NADPH, etc. [20]. Therefore, OM transporters typically require an inner membrane bound protein complex to facilitate nutrient transport [21, 22, 23]. One such example is the TonB-dependent transporter (TBDT) family of OM proteins, which require the inner membrane bound TonB:ExbB/D complex to facilitate the import of metal-containing complexes into the periplasm [23]. The N-terminus of TonB and ExbB/D span the IM and transduce energy from the proton motive force, generated by the electrochemical gradient across the IM, to extract the luminal domain of TBDTs. It is still not clear how TonB can transduce this energy to TBDTs in the OM nor conversely how TBDTs signal TonB to extract their luminal domain.

A key feature of TBDTs is that they import large, metal containing complexes, e.g. ferric siderophores and chelates, cobalamin, zinc, etc. across the OM [23]; such that TBDTs have evolved various techniques to pirate metal complexes from their host organism, as is the case with Transferrin binding protein (TbpA/B) in *Neisseria gonorrhoeae*. Because of their molecule specificity and their ability to transport large nutrients, the TBDT family makes an ideal drug target [24]. Furthermore, because metallic compounds are required for metabolic processes inside the cell, hindering the function of TBDTs can inhibit enzymatic processes, thereby slowing bacterial replication or proliferation, and potentially increase the efficacy of antibiotics. Structurally, TBDTs are characterized by a 22-stranded $\beta-$barrel motif with a luminal, or plug, domain that occludes the barrel [23]. At the N-terminus of the luminal domain is a highly conserved sequence of amino acids known as the Ton box, which is where TonB binds to facilitate transport [25]. It has been shown experimentally that deletions or mutations to the Ton box can drastically reduce nutrient import through specific TBDTs [26]. Furthermore, as is the case for the vitamin B12 uptake protein (BtuB), some TBDTs require ions or, as is the case for TbpA, assistant proteins

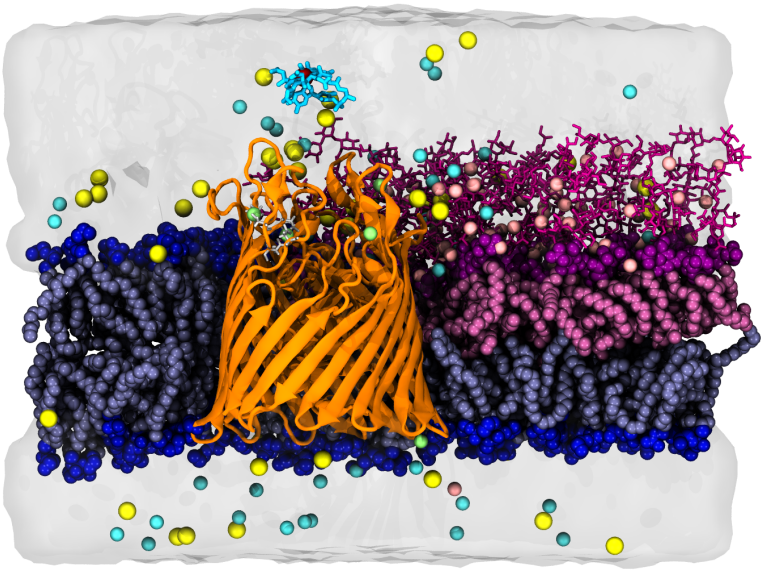


Figure 1.3: A) All-atom representation of the vitamin B12 transport protein BtuB (orange ribbons) in a symmetric (left) and asymmetric (right) bilayer. The LPS oligosaccharides reduce extracellular loop fluctuations by excluding the volume around the protein. Lipid tails and heads are shown in glossy spheres with the LPS oligosaccharides shown in magenta sticks. B12 is shown in light blue sticks with the cobalt shown as a red sphere. The calcium binding site for BtuB is shown in white sticks and ions are represented as spheres.

for the binding of substrate [27, 2, 28]. Therefore, understanding the strength of substrate binding as well as the flexibility of the binding site could illuminate the method by which TBDTs signal TonB, how the luminal domain unlocks to permit substrate transport, and how the proton motive force from TonB is transduced to the TBDT.

1.3 Utilizing Mass Repartitioning to Accelerate Lipid Simulations

Membrane models are frequently used in molecular dynamics (MD) studies because of their biological relevance. As such, MD simulations can be used to study membrane properties and provide an atomistic description of membrane structure and dynamics [29, 30, 31, 32, 33, 34, 35]. For example, membrane-embedded proteins often require the incorporation of a membrane in the model for protein stability [36, 37, 38]. Additionally, membrane permeability and small molecule interactions are often of interest in drug design, which can be investigated computationally using membrane models. [39, 40]

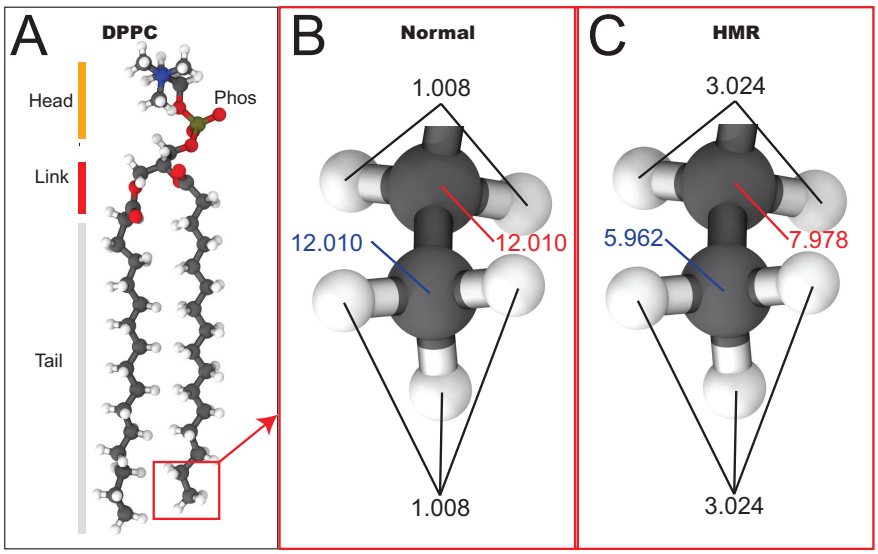


Figure 1.4: (A) Structure of DPPC lipid, phos and link denote phosphate and ester link, respectively. (B) Normal mass distribution for the atoms in the acyl-chain tail. (C) Mass distribution for the same atoms with HMR.

In order to make predictions from MD simulation, accurate lipid force fields are necessary when studying membrane systems[41, 42]. To date, multiple force field models are available: AMBER14 [43], SLIPIDS [44], CHARMM36 [45], and multiple GROMACS united atoms models [46, 47, 48]; several studies have compared present lipid models and improved upon existing ones[39, 29, 48, 41, 42]. Force-field parameters are typically evaluated based on their ability to reproduce experimentally known structural and dynamical properties of pure-lipid bilayers, e. g., lipid area, bilayer thickness, compressibility modulus, deuterium order-parameters, and diffusion coefficients[29, 42, 35, 41]. In some cases partition coefficients have been calculated to validate lipid:small-molecule interactions [34]. Furthermore, compatibility with water and protein force fields should also be considered when choosing a lipid force field[49, 50].

The CHARMM36 (C36) lipid force field is frequently used in MD simulations because it can accurately reproduce a number of physical properties of lipids, as well as its compatibility with the C36 protein and general small-molecule force field [45, 29, 42, 34, 35, 49, 50]. The most recent C36 lipid force field update resulted in improved

agreement with experimental order parameters, compressibility modulus, and area per lipid [45]. Furthermore, the recently launched CHARMM-GUI web interface, which supports several MD software packages, has greatly facilitated the construction of membrane systems for MD simulations, specifically utilizing the C36 force field [51, 52, 53, 54]. CHARMM-GUI automatically generates structure, coordinate, parameter and input for pure membrane and membrane-protein systems, supporting numerous phospholipid molecules through lightweight object oriented structure–analysis (LOOS) [55, 56].

In order to study properties of a membrane system, multiple simulations of sufficiently long time scales are typically required [57, 39]. As such, there is significant benefit to running longer MD simulations without the prohibitive computational cost, particularly for membrane-containing systems, which can be much larger than protein-only systems. One previously suggested approach to speed up MD simulations is by using a longer time step [58, 59]. Implementing a longer time step decreases the accuracy when integrating the equations of motion in MD; however, it has been shown that the introduced errors are typically much smaller than the statistical errors due to limited sampling. [60] Additionally, the increased energy drift introduced by a longer time step can be dampened by using a thermostat [60, 59]. Currently, the magnitude of the time step in atomistic MD is limited by the fastest moving atoms in the simulation, which are the vibrational motions of the hydrogen atoms [58]. Therefore, increases in time step can be achieved by slowing down or restricting the movement of the hydrogens [58]. Common practice in MD simulation has been an increased time step from 1 to 2 fs by keeping the covalent bonds involving hydrogen atoms rigid using SETTLE and SHAKE algorithms for water and other molecules, respectively [61, 62]. The implementation of these algorithms doubles the available simulation time at fixed computational cost; however, the SHAKE algorithm is likely to fail at time steps beyond 2 fs for conventional MD [61, 59].

Recently, MD simulation studies have shown that time steps of up to 4 or 5 fs can be altering hydrogen masses [58, 59, 63]. It is important to note that when implementing

mass modifications, it is requisite that the total mass of the system does not change [58, 59]. As described by Feenstra et al., increasing the the total mass of the system will result in a slower time scale for various events of interest, e. g., diffusion [58]. Similarly, in the virtual site technique (VST), the hydrogens' masses are assigned to the adjacent heavy atoms and their positions are calculated and updated based upon the positions of the heavy atoms [64, 63]. However, implementing VST requires re-optimization of force field parameters, such that when applied to the C36 force field, VST was shown to alter several lipid properties, leading to thinner and more disordered bilayers [63]. Recently, it was shown that the combination of VST with HMR on every fourth methyl group in the lipid tails resulted in strong agreement with measured lipid properties in standard MD simulations [65].

To clarify, HMR modifies the atomic input by repartitioning mass from each heavy atom to its covalently bonded hydrogens, while conserving the overall molecular mass [58, 59]. For example, when applying HMR, the reweighted hydrogen mass should not be greater than 3 amu because a larger mass transfer would make methyl carbons lighter than their bonded hydrogens. Hopkins et al. illustrated this method for both a small peptide as well as a large protein in explicit solvent [59]. They found that the protein/peptide with HMR applied, using a hydrogen mass of 3 amu and a time step of 4 fs, consistently reproduced conformations observed without HMR. However, when HMR was applied to both protein and water molecules, there was an increase in the viscosity of water and, consequently, slower transition rates between different protein conformations. Therefore, HMR should not be applied to water. Since its inception, HMR has been used in several software packages such as NAMD [66], AMBER [67] and ACEMD [68], in order to speed up MD simulation output.

It has also been shown that additional speed up can be obtained by decreasing the cutoff for non-bonded interactions. Although C36 lipids were parameterized and validated using a 12-Å cutoff with a switching function applied at 10 Å [45], most HMR studies employing lipids to date have used the C36 lipid force field with a 9 Å (default setting) cutoff in

ACEMD or AMBER [69, 70, 71]. Previous MD simulations have shown that membrane properties are highly sensitive to the cutoff value and Lennard-Jones (LJ) switching functions because lipids dynamics are primarily drive by LJ interactions [45, 54]. Additionally, lipids are more hydrogen rich than proteins, for which HMR was previously validated [59]. However, to date, the effects of HMR and a shorter 9-Å cutoff with the C36 lipid force field have not been investigated.

1.4 Towards a Cell Envelope Model

Recall that the cell envelope of Gram-negative bacteria is a multi-component system consisting of a symmetric IM, a peptide cross-linked cell-wall, and an asymmetric OM. Each membrane, IM and OM, have an array of integral membrane proteins and associated lipoproteins that have various forms and serve multiple functions. Much work has been done at the cell IM, and results from the SimBac lab have elucidated the action of the cell-wall [72, 73, 74]. Furthermore, results presented within this thesis [75] as well work performed in the OM [76, 77] have shown the importance of simulating the native or natural environment. Therefore, building upon previous molecular dynamics simulations, protein abundance data, whole cell genomics, and currently known protein structures, we present a method for constructing an all-atom cell envelope for Gram-negative bacteria, which can be further implemented within other Gram-negative bacteria. Furthermore, we suggest more analysis that can be performed provided input data to create various cell-envelope microcosms with the associated proteins in appropriate abundances for study of large-scale, envelope-spanning, protein complexes.

CHAPTER 2

ROLE OF THE NATIVE OUTER-MEMBRANE ENVIRONMENT ON THE TRANSPORTER BTUB

2.1 Introduction

TonB-dependent transporters, a subset of OM proteins, are 22-stranded, monomeric β-barrels that transport large and rare nutrients, typically ferric siderophores and chelates, by coupling to an energy-transducing, cytoplasmic-membrane-bound TonB complex that spans the periplasm [78, 79, 23, 80]. While there are many similarities among TonB-dependent transporters, BtuB uniquely depends on calcium ion binding to transport cobalamin (CBL), also known as vitamin B_{12}, across the OM [27, 2, 28]. Crystal structures show that five aspartate residues on EC loops 3/4 and 5/6, constructing the so-called "aspartate cage" (Asp-cage), provide the binding site for two Ca^{2+} ions necessary for CBL transport [17, 18]. Loop positions are known to be sensitive to solutes [81] and protein dynamics are sensitive to membrane composition and thickness [82], necessitating highly accurate models to capture such sensitivity.

In this work, we first introduce an LPS-containing model of the OM constructed *in silico* in order to perform MD simulations of OM proteins in their native environment. We then calculate the area per lipid, hydrophobic thickness and lateral diffusion rate of the asymmetric bilayer. Next, we use this OM model to study the interactions of BtuB with LPS molecules and compare it to those of BtuB in a symmetric phospholipid bilayer, revealing unique protein-LPS interactions in the asymmetric system. Finally, this model of

the OM is used to elucidate the role of Ca^{2+}-binding, which is required for high-affinity binding of CBL [2, 28].

2.1.1 Experimental Procedures

Simulation Systems

Five systems were built containing BtuB to elucidate the effect of LPS and the role of Ca^{2+}-binding in EC-loop stability. The apo-state (PDB: 1NQE) structure was placed in a symmetric 1-palmitoyl 2-oleoyl-phosphotidyl-ethanolamine (POPE) bilayer and two asymmetric LPS (outer leaflet) - POPE (inner leaflet) outer membranes using VMD [83, 84]. The Ca^{2+}-bound (PDB: 1NQG) state was also placed in two asymmetric model membranes [83]. The two asymmetric outer membranes only differ in the initial arrangement of LPS molecules around the protein. Missing extracellular loops in the crystallographic structures of BtuB were completed using the TonB-bound structure (PDB: 2GSK) [25, 18], as seen in Figs. 2.1A and 2.1B.

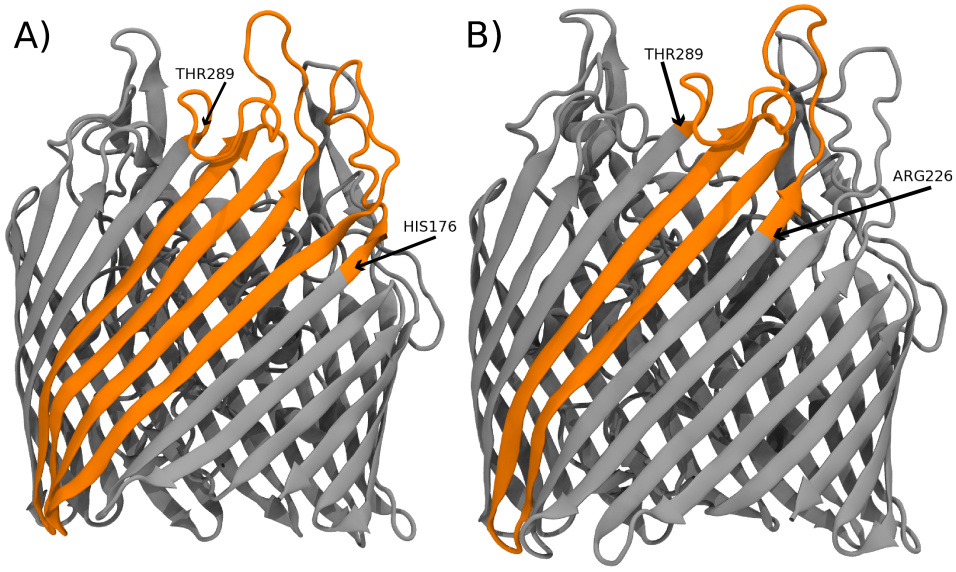


Figure 2.1: A graphical representation of the loop completions performed on the A) apo-state BtuB (PDB:1NQE) and the B) Ca^{2+}-bound BtuB. Loop completions, shown in orange, utilized the TonB-bound BtuB structure (PDB: 2GSK).

Each system was solvated and then ionized to a salt concentration of 150mM NaCl. The apo symmetric (Apo-sym) system has a total of ∼70,000 atoms with an outer leaflet composed of 66 POPE lipids and an inner leaflet composed of 61 POPE lipids. The disproportionate number of lipids is due to the cross sectional area of BtuB being greater on the periplasmic side than on the extracellular side; therefore, more lipids were needed in the outer leaflet to keep the area per lipid roughly identical between layers. The asymmetric-bilayer systems all have a total of ∼137,000 atoms with an outer leaflet composed of 50 LPS molecules and an inner leaflet of 157 POPE lipids, details can be found in Table 2.1.

Table 2.1: Summary of simulated systems. Lipid counts are reported in Table 6.1.

System	Atoms	Water molecules	Ions
Apo-sym	69494	14823	99
Apo1-OM	137411	27019	452
Apo2-OM	137407	27019	448
Ca1-OM	137407	27019	448
Ca2-OM	137407	27019	448
OM only	71781	13068	207

The initial model used for a single LPS was taken from the crystal structure of LPS in complex with FhuA [85]. This LPS from *E. coli* K-12 is the rough form, also known as the RaLPS chemotype [4], containing only lipid A along with the 10-saccharide core region (see Fig. 2.2 for a schematic of the core saccharide) [86]. Two phosphate groups were added to heptose I and one group to heptose II. Based on the crystal structure occupancy [87], half of the lipid A moieties have one phosphate on the first glucosamine and half have two; all have one phosphate on the second glucosamine. Thus, the net charge on each LPS molecule is either -11e or -12e. LPS molecules were neutralized using mostly (97%) Mg^{2+} ions, along with a few (3%) Ca^{2+} ions. These ions were initially placed using the cIonize plugin in VMD, which uses a Poisson-Boltzmann solver to iteratively place ions in electrostatic minima [88]. Force-field parameters for LPS were constructed by analogy using the CHARMM36 lipid and carbohydrate force fields and, thus, are nearly identical to those developed in Wu et al. (see Supplementary Materials for details) [9].

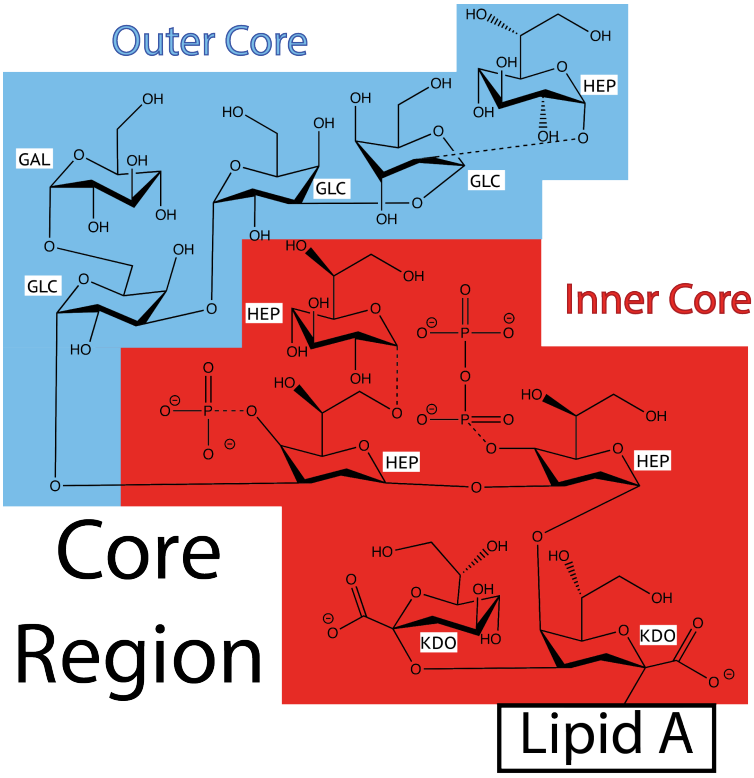


Figure 2.2: Chemical structure of the core oligosaccharide used in the OM model membrane simulations. Chemical structure is build in regards to the K12 *E. coli* oligosaccharide outlined in Nikaido 2003.

Molecular Dynamics Protocol

NAMD2.9 was used for all MD simulations [66] along with the CHARMM36 all-atom force field for protein, ions and phospholipids [50, 45]. A 2-fs time step was employed. Bonded terms were evaluated every time step and non-bonded terms and long-range electrostatic interactions were updated every 2 fs and 4 fs, respectively. Long-range electrostatic interactions were calculated using the particle-mesh Ewald (PME) method [89]. Lennard-Jones long-range interactions were cutoff at 12 Å with a force-based switching function used beginning at 10 Å. Temperature and pressure were maintained at 310K and 1 atm through Langevin dynamics and a Langevin piston, respectively. All bonds involving hydrogen atoms were constrained using the SHAKE algorithm [61].

System equilibration was carried out over multiple steps. First, all atoms were constrained with the exception of lipid tails, which were equilibrated for 1 ns allowing them to melt. Second, water, ions, and lipid head groups were released and equilibrated for 2 ns in the symmetric bilayer. Due to large initial spacing of the LPS molecules, which was necessary to avoid clashes during building, a surface tension of -100, -200, -250 and -300 dynes/cm equilibrated at 2 ns each was applied to the asymmetric membrane. This process allowed the LPS to eliminate any gaps introduced in construction. Protein side chains were then released for 1 ns before releasing the protein backbone and equilibrating for 10 ns. Simulations were run in the NPT ensemble for 150 ns for the Apo-sym and 300 ns for the Apo1-, Apo2-, Ca1- and Ca2-OM systems after equilibration. Analysis of each system was performed on data collected from the last 200 ns of production for the OM systems and the last 100 ns of production for the symetric bilayer, as shown in Fig. 6.1. Langevin dynamics were employed to maintain the temperature at 310 K and a Langevin piston was used to keep the pressure at 1 atm throughout equilibration and production simulations.

For steered molecular dynamics (SMD), an imaginary particle was affixed to the acetylated N-terminus of BtuB. This imaginary particle was pulled at constant velocities of 1 Å/ns and 0.25Å/ns normal to the bilayer plane out of the barrel. Twenty-two C_α atoms in a ring around the barrel, one per beta-strand, were fixed normal to the bilayer to prevent system propagation due to luminal domain extraction. The 1Å/ns SMD runs were performed three times for each Apo system and the results presented are system and time averaged. Due to the amount of time required, the 0.25Å/ns SMD runs were only performed once each for the Apo-sym and Apo-OM.

Analysis

The membrane thickness was calculated by measuring the average distance between acyl chain C2 atoms in the POPE leaflet and C2 or C4 on each of the LPS tails, as done previously [90, 53]. The APL was calculated by determining the total lateral surface area of the

system, subtracting off the cross sectional contribution from the protein, and then dividing that area by the number of lipids in each leaflet. The cross sectional area of BtuB was determined using a double summation technique for a plane width of 1 Å normal to the axis of the protein between the periplasmic turns and extracellular loops. Lateral diffusion coefficients were obtained by calculating the mean-square displacement of the POPE hydrophilic head center of mass as done elsewhere [91, 92]. Similarly, diffusion coefficients for LPS were calculated by measuring the mean-square displacement of the lipid A center of mass. Mean-square displacement was also determined for the LPS center of mass, which resulted in nearly identical diffusion values (data not shown). RMSF values were calculated only for the C_α atom of each amino acid [93] to remove rotamer conformation bias. A full summary of the aforementioned values for each system can be found in Table 6.1.

2.2 Results and Discussion

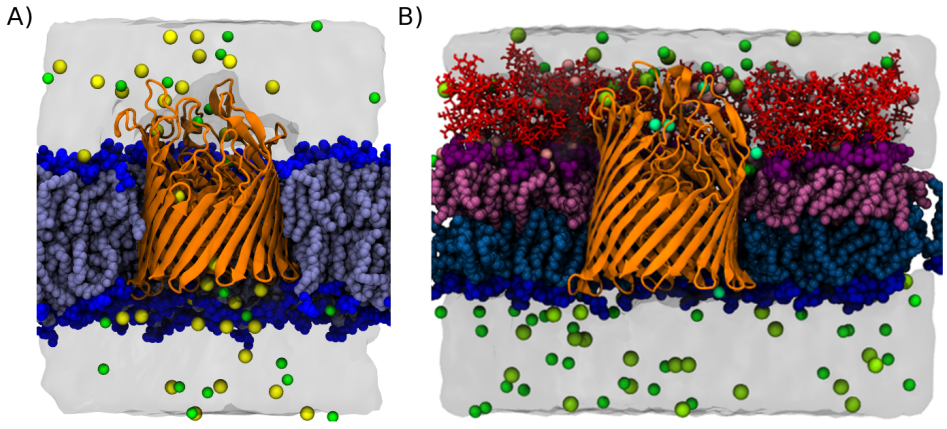


Figure 2.3: Snapshot of BtuB in (A) a symmetric bilayer and (B) an asymmetric bilayer. Water is shown as a molecular surface. Ions are shown as yellow for Na^+, green for Cl^-, pink for Mg^{2+}, and teal for Ca^{2+}. NaCl concentrations are 150mM in each system. BtuB is represented as orange ribbons. The phospholipid tails are shown in light blue and head groups as dark blue spheres. Lastly, Lipid A of LPS is shown in pink and corresponding Lipid A head groups as magenta spheres; the core oligosaccharides are shown in red sticks.

2.2.1 Bilayer dynamics

Although the OM typically has an inner leaflet composition similar to the cytoplasmic membrane [94, 95], the OM model used in this study is simplified by utilizing an inner leaflet composition of pure POPE phospholipids. This is an intuitive simplification since the OM inner leaflet is composed of $\sim$80% PE lipids [94]. We model our OM after K12 *E. coli*, whose LPS lacks O-antigen chains, possessing only the membrane-forming lipid A and the inner and outer core oligosaccharides [6, 13, 10]. The model OM is shown in Fig. 2.3 with the symmetric bilayer for comparison. In order to compare our model system with previous experimental and computational results, we calculate the hydrophobic thickness, lateral diffusion, and area per lipid for the apo BtuB symmetric bilayer system (Apo-sym), Fig. 2.3A, each apo BtuB asymmetric bilayer (Apo1-OM, Apo2-OM), and each Ca^{2+}-bound BtuB asymmetric bilayer (Ca1-OM and Ca2-OM) systems, Fig. 2.3B. We also simulated the model OM with no protein present.

The first check of model validity is the hydrophobic thickness of each membrane system. As shown in Fig. 6.2, each protein-containing OM system as well as the model OM has a hydrophobic thickness of approximately 25 Å, which matches well with recent results of another OM simulation model [96]. When comparing the OM systems to the Apo-sym profile shown in Fig. 6.2A, it is clear that the asymmetric OM is much thinner than the symmetric bilayer. The membrane thickness difference is attributed primarily to tail length difference (LPS has 12 to 14 carbons per tail whereas POPE has 16 to 18) as well as lipid tail inter-digitation in the OM systems. Another verification of the hydrophobic thickness is the location of aromatic residues on the protein, specifically phenylalanine, tryptophan, and tyrosine, which are known to reside at amphiphilic interfaces [97, 98]. From the density profiles, seen in Fig. 6.2, it is clear that the OM systems match well with the aromatic residue profile of the protein. In contrast, the increased hydrophobic thickness of the Apo-sym system permits PE head groups to interact with the protein along the β-barrel/EC loop interface, which can be seen in Fig. 6.3. However, the oligosaccharide number density as

shown in Fig. 6.2B-E indicates that the LPS head groups align with the EC loops of the protein.

The next check on model validity is the average lateral diffusion of LPS in the OM systems. It is understood that ionic bridging between core oligosaccharides by divalent cations significantly decreases lateral mobility of LPS [6]. Multiple fluorescence experiments provide a range on LPS lateral diffusion from 10^{-13} to 10^{-9} cm^2/s depending upon the LPS variant [11, 99, 100, 101], This is in stark contrast to phospholipid diffusion values, which are on the order of 10^{-8} cm^2/s [45]. Shown in Fig. 6.4 and elaborated in Table 6.1, the diffusion coefficients determined for LPS in the OM systems vary between 6.0 and 9.8 $\times 10^{-10}$ cm^2/s, which are comparable to values found in previous simulations [8, 9]. Diffusion coefficients for the inner leaflet of phospholipids are much larger, ranging between 4.3 and 6.4 $\times 10^{-8}$ cm^2/s, for all protein-membrane systems simulated. This range is only slightly higher than the calculated diffusion coefficient for the Apo-sym bilayer, namely 2.0 $\times 10^{-8}$ cm^2/s.

The last check on the OM model is the observed area per lipid for each system. The pure OM simulation was initially performed with 36 LPS and 102 POPE. This ratio was initially chosen to replicate an experimental value of $\sim$156 Å^2/LPS [102], which was utilized in simulations by the Khalid group [8]. The area per LPS was found to be 179 Å^2/LPS for this test system, which was much closer to that of a similar variant studied by Wu et al. [9]; a full comparison can be found in Table 6.1. This value further contrasts with the measured area per LPS for the protein-containing OM systems, which was $\sim$10% larger, between 195 and 199 Å^2/LPS. This is partly because the area per lipid or LPS for protein-containing bilayers is less reliable than for pure bilayers as the cross-section of the protein varies over its profile, making it difficult to estimate a priori the number of lipids in each leaflet prior to simulation. While the area per LPS differs in the pure-OM and protein-OM systems, the large number of LPS-protein interactions (see Tables 6.2 and 6.3) and the consistent gel-like nature of the LPS leaflet suggest that this small difference in density does not lead

to a substantive difference in BtuB's behavior.

A simulation of a pure POPE bilayer with 102 lipids per leaflet was also carried out for 60 ns. After disregarding the first 10 ns, the area per lipid was found to be 57.9±0.9Å^2, which is near the CHARMM36 reported area per lipid for POPE of 59.2±0.3Å^2 [45]. This value also is only slightly less than the measured area per lipid in the Apo-sym case of 58.9 Å^2. The area per lipid for the inner leaflet in the OM systems was observed to be 4-8% higher at 61.4-63.7Å^2 (see Fig. 6.5 and Table 2.1), although this may be ascribable to inaccuracies in protein area. In light of the results from Wu et al. and the results of the simulations reported here, future systems should be adjusted to target an area per LPS of 180 Å^2.

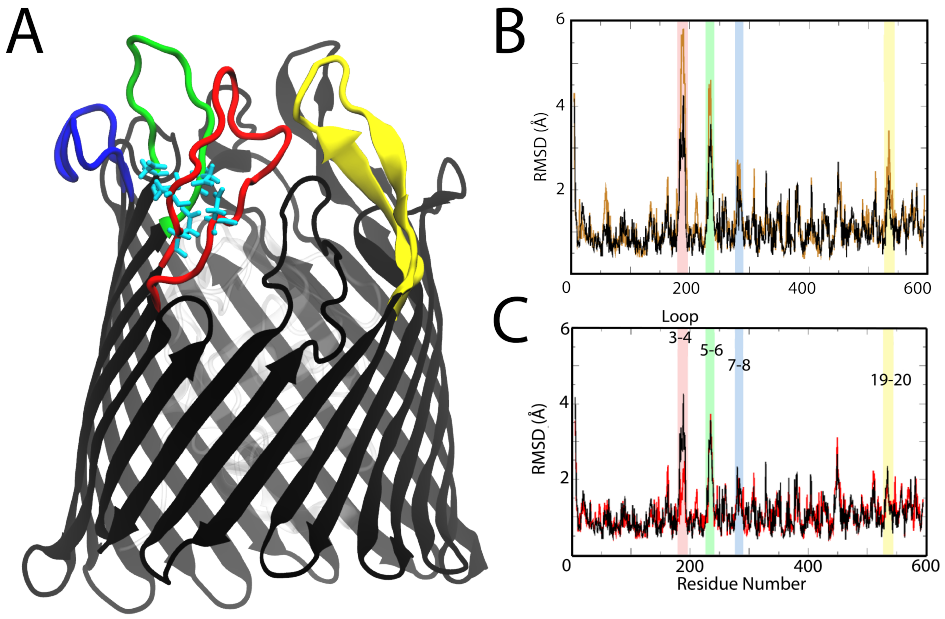


Figure 2.4: (A) Ribbon representation of BtuB with Asp-cage highlighted in cyan sticks. Key extracellular loops highlghted in red (loop 3/4), green (loop 5/6), blue (loop 7/8), and yellow (loop 19/20). (B) Plot of root mean-square fluctuations (RMSF) by residue of the BtuB Apo-sym (orange trace) and Apo1-OM (black trace) bilayers. Substantial RMSF reduction occurs in the highlighted regions, which correspond to the highlighted loops in (A). (C) Plot of RMSF by residue for the Apo1-OM (black trace) and Ca1-OM (red trace) systems. A slight difference in RMSF is located in loop 3/4, where Ca^{2+} binding occurs. RMSF plots for Apo2-OM and Ca2-OM can be found in Fig. 6.6.

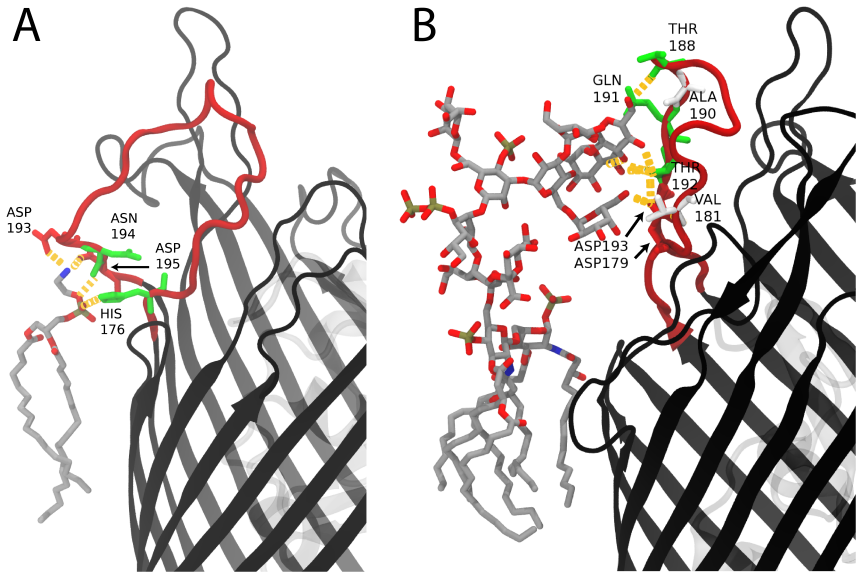


Figure 2.5: Hydrogen bonding interactions between the protein and membrane. Specific protein-lipid hydrogen bonds are indicated as orange dashed lines and BtuB is shown in black ribbon. (A) Apo-sym system showing a phospholipid, shown as grey sticks, hydrogen bonding (orange dash) to the loop 3/4 of BtuB. Loop 3/4 becomes deformed as Asp193 and Asp195 are attracted toward the zwitterionic POPE head group. This interaction causes loop 3/4 to substantially deviate from its conformation observed in the CBL-bound structure. (B) Apo2-OM system with an LPS molecule, shown in grey sticks, hydrogen bonding (orange dashes) to loop 3/4 of BtuB. The oligosaccharides in the OM systems exhibit hydrogen bonding along the extracellular loops of, including loop 3/4 residues Asp179 and Asp193, in the absence of Ca^{2+}. Further hydrogen bonding interactions can be found in Tables 6.2, 6.3.

2.2.2 Apo-state comparison

To better understand how LPS affects BtuB function, we first compared the interactions of the apo-state BtuB, i.e., no Ca^{2+} nor CBL, in the Apo-sym and Apo-OM systems described above. The asymmetric distribution of outward-facing charged and polar amino acids on OM proteins [97], substantial anionic charge of LPS, and the observed alignment of the oligosaccharides with EC loops suggest LPS should affect protein functionality. This claim is substantiated by observing the stabilizing role LPS has on the EC loops, organizational change in the interior luminal domain of BtuB, and significant difference in force required to partially extract the luminal domain.

The stabilizing role of LPS on the protein is exemplified by the root mean-square fluctuation (RMSF) of the protein backbone, particularly EC loops, shown in Fig. 2.4B. Each of the highlighted regions, which correspond to highlighted loops of BtuB in Fig. 2.4A, shows a 1-2 Å stabilization when LPS is present. This effect is due to the alignment of the oligosaccharide along the EC loops as seen in Fig. 2.5B, which prevents the loops from folding back over the membrane and, thus, limiting their range of motion, at least on the $\sim$300-ns time scales investigated here. Further investigation of the oligosaccharide-protein interaction reveals a dense hydrogen bond network between the LPS and EC loops. These hydrogen bonds occur between the LPS and charged, outward-facing residues in the EC loops, particularly those closer to the EC loop apex shown in Fig. 6.3 and enumerated in Table 6.2. This contrasts with the membrane-protein interactions observed in the Apo-sym system, which occur proximal to the hydrophobic-matching region of the protein, along the barrel/EC loop interface (see Fig. 2.5A).

An unexpected effect observed in the Apo1- and Apo2-OM simulations is a conformational change in the luminal domain of BtuB. The first observation was a structural difference between the symmetric and OM systems in the region from Gly82 to Ser91. This region forms an α-helix in the Apo-sym simulation within the first 10 ns, whereas in all OM systems this region forms a random coil, as seen in Fig. 2.6. Comparing this α-helix in

the Apo-sym system to the CBL-bound crystal structure (PDB: 1NQH) reveals that Ser91, a CBL-binding residue, is farther from its CBL-bound-state position due to this helix. In contrast, the Apo1- and Apo2-OM simulations exhibit excellent alignment with the Gly82 to Ser91 backbone and side-chain orientations found in the CBL-bound crystal structure within the first 20 ns of simulation. These structural changes observed are then maintained throughout the remainder of each simulation (see Fig. 6.7). The second observation was a conformational difference of a side chain rotamer in the Apo-sym and Apo-OM simulations. Arg14, which forms a hydrogen bonding network with Gln299, Asp316, and Arg358 in the Apo-sym simulation, rotates about its C_β atom in the Apo1-OM toward Gly82. This rotation effectively "unlocks" the so-called Ton-box, the N-terminal residues that mediate interaction with TonB, from the barrel wall. An unlocked Ton-box is also seen in the CBL-bound crystal structure, although in neither case is full release into the periplasm observed.

To determine the functional role of the observed structural difference in the luminal domain and the apparent unlocking of Arg14, steered molecular dynamics (SMD) simulations were performed to extract the Ton-box in the Apo-sym, Apo1- and Apo2-OM systems at a rate of 1Å/ns and Apo-sym and Apo1-OM at 0.25Å/ns. Although exceedingly fast compared to the biological process, these rates are in line with previous simulations [1]. It was found that the unlocking of Arg14 results in an approximately 30% lower force required for Ton-box extraction for the apo-OM systems compared to the Apo-sym system at 1Å/ns(see Fig. 2.7). The approximately 250-pN higher force peak between 9 and 12 ns in the Apo-sym case is due to Arg14 being electrostatically bound to the barrel wall via the hydrogen bond network described above. The 500-pN force around 5 ns in both averages correlates with the unbinding of Val10, Thr11, Ala12 from the luminal domain hydrophobic core. This force would be significantly reduced upon TonB binding due to favorable hydrophobic interactions between the Ton box and the C-terminus of monomeric TonB [25, 80]. The observed difference is further confirmed at 0.25Å/ns, which shows an approximately 15% lower force required for Ton-box extraction in the OM system.

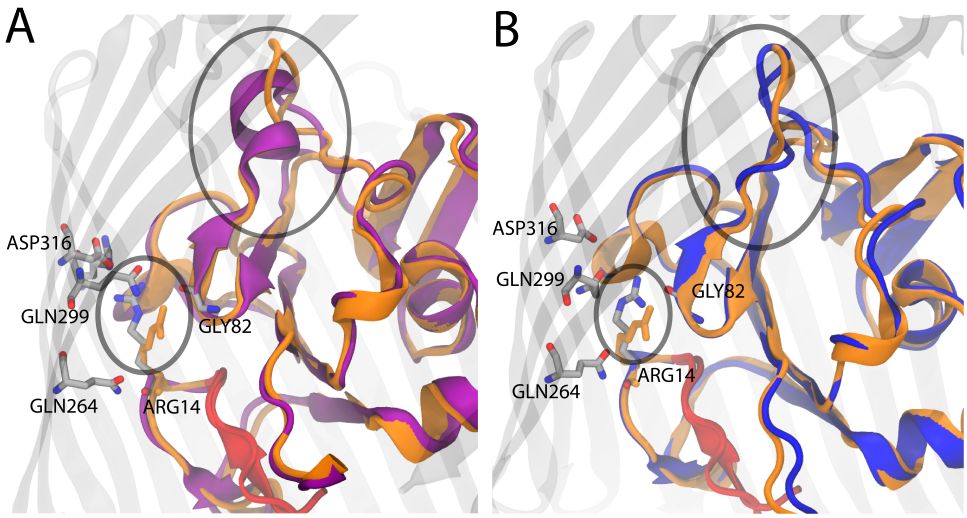


Figure 2.6: Luminal domain organization of the (A) Apo-sym (purple ribbon) and (B) Apo-OM (blue ribbon) systems compared to the CBL-bound crystal structure luminal domain (orange ribbon). There is an observed difference in the luminal domain secondary structure (large oval) of BtuB in the simulated membrane systems. Apo-sym reveals an $\alpha-$helix between Gly82 and Gly92 whereas each of the OM systems reveals a random coil along the same residues (see also Fig. 6.7). The random coil conformation brings Ser91, a CBL-interacting residue, closer to the its conformation observed in the CBL-bound crystal structure. The small oval shows the Arg14 (grey sticks) lock mechanism (A) "locked" to the barrel wall, as found in the symmetric system and (B) "unlocked" or bend toward the lumen. The "unlocked" conformation of Arg14 observed in the simulated OM systems is also found in the CBL-bound crystal structure (orange stick). Movies of lumen extraction using SMD can be found through Biophysical Journal.

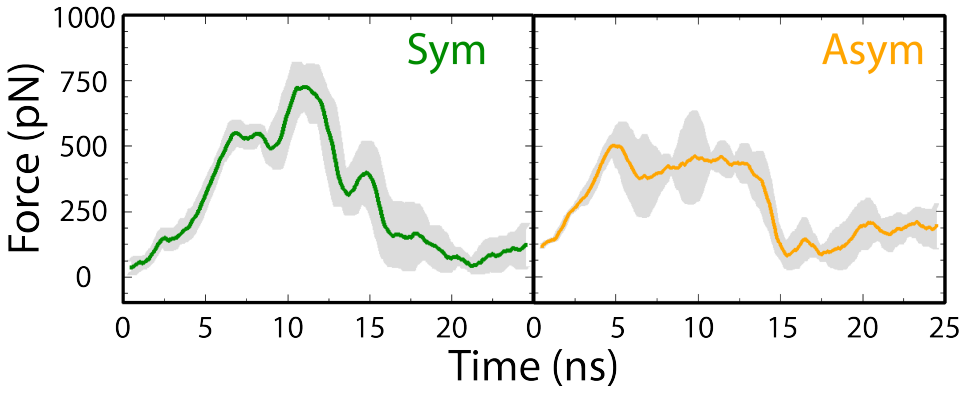


Figure 2.7: Plot of the observed force for Ton-box extraction from the luminal domain in (A) Apo-sym (green) and (B) Apo-OM (orange) systems with corresponding standard error shown in grey. Each curve is an average of three time-averaged constant velocity SMD runs of 1 Å/ns. The Apo-OM curve is the run average of 2 Apo1-OM and 1 Apo2-OM SMD runs. The peak force for Apo-sym occurs around 11 ns and corresponds to unbinding of Arg14 from the hydrogen bond network as seen in Fig. 2.6. Peak force in Apo-OM systems occurs around 6 ns and corresponds to the hydrophobic interaction of Val10 to Ala12 with the luminal domain. Independent SMD run results can be found in Fig. 6.8.

2.2.3 Calcium-bound state comparison

Previous experiments by Bradbeer et al. concluded that BtuB has a significantly higher binding affinity for CBL in the presence of calcium ions [2, 28]. Furthermore, previous MD simulations by Gumbart et al. showed that calcium binding decreases the fluctuations of the CBL-binding residues compared with the apo state, suggesting that the role of calcium is to stabilize the binding site [18]. However, upon comparing the Apo-sym and Apo-OM systems, we observe that LPS significantly stabilizes many of the CBL-binding residues located on the EC loops as seen in Fig. 2.4B. Thus, the importance of Ca^{2+} binding in the Asp-cage remains unclear.

Five calcium ions are found in the crystallographic, CBL-bound structure of BtuB, two of which are bound to the Asp-cage located on EC loops 3/4 and 5/6 [17, 18]. Earlier mutagenesis experiments on residues in the Asp-cage revealed a dependence on Asp residues on loop 3/4 for effective CBL uptake [28]. To evaluate the importance of calcium, we again compare the RMSF of the Ca-OM systems to the Apo-OM systems to illuminate a potential

role for Ca^{2+}-binding in the Asp-cage. Fig. 2.4C and S8, shows a substantial reduction of RMSF in loop 3/4 in the Ca-OM systems, the RMSD is shown in Fig. 6.10; however, Ca^{2+} has little to no additional stabilizing effect on the other EC loops. Furthermore, the luminal domain takes on a similar conformation as observed in the Apo-OM systems, as shown in Fig. 6.7, indicating that the OM system promotes the proper luminal domain organization. Therefore, combining observations from the present simulations with previous experimental evidence indicates that Ca^{2+} binding primarily affects loop 3/4. However, high-affinity binding of substrate is not dictated solely by protein fluctuations near the binding site and, thus, stabilization of binding-site loops alone may be insufficient. The presence of calcium in the Asp-cage induces loop 3/4 into a conformation that closely resembles the CBL-bound structure of BtuB, which was not seen in either of the Apo-OM simulations. As shown in Fig. 2.8A, most of the CBL interacting residues on the EC loops align with their observed conformation in the CBL-bound crystal structure; however, Asn185 on loop 3/4 is poorly aligned. Comparatively, with Ca^{2+} bound as shown in Fig. 2.8B, there is a nearly perfect alignment of Asn185 to the CBL-bound crystal structure, which, after forming, is maintained for the remainder of the simulations (see also Fig. 6.7). Thus, Ca^{2+} binding not only promotes stabilization but also the structural reorganization of loop 3/4.

The arrangement of residues in the Asp-cage on loop 3/4 due to Ca^{2+} binding induces a conformational shift of the remaining loop residues. This shift is made apparent by the spontaneous formation of a small α-helix between Tyr183 and Gly187, seen in Fig. 2.8B. The α-helix is first observed 7 ns into the Ca1-OM trajectory and is maintained for aprroximately 50% of the trajectory thereafter. The helix is not formed in the Ca2-OM trajectory; however, loop 3/4 still aligns well with the CBL-bound crystal structure as seen in Fig. 6.9. Other recent molecular dynamics simulations have also observed spontaneous helix formation for an OM protein, OmpLA, in a native OM bilayer [53].

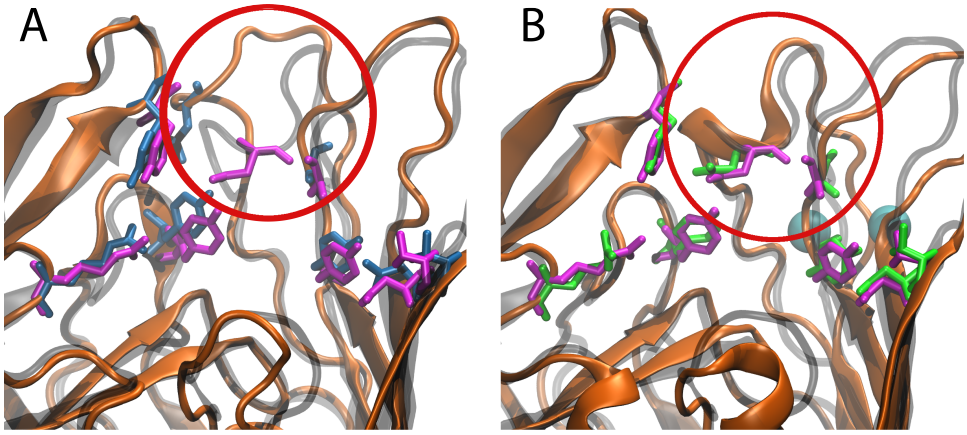


Figure 2.8: Comparison of the (A) Apo- and (B) Ca^{2+}-bound states of BtuB in the simulated OM systems. Orientation differences observed in loop 3/4 are emphasized by the red circle. (A) Apo-state BtuB (orange ribbon) aligned to the CBL-bound crystal structure (transparent grey ribbon) from Chimento et al. [83]. Residues within 4 Å of CBL in the CBL-bound crystal structure are highlighted as blue (Apo1-OM) and purple (crystal structure) sticks. (B) Ca^{2+}-bound BtuB (orange ribbon) aligned to the CBL-bound crystal structure (transparent grey ribbon). Calcium ions are shown as cyan spheres and the green sticks represent the CBL binding residues in the Ca1-OM system. Replicates from the second OM system can be seen in Fig. 6.9.

2.3 Conclusions

The outer membrane of Gram-negative bacteria provides an additional barrier to nutrients, antibiotics, pathogens and other large molecules. This barrier provides the bacterium with added protection but conversely makes uptake of essential nutrients more difficult. To better understand the local effects of the outer membrane on embedded proteins, we developed an OM model suitable for MD simulations. The hydrophobic thickness, area per lipid and lateral diffusion rates of LPS in this OM model correspond reasonably well with previous simulations and experiments [53, 9, 93, 90, 15]. The small discrepancies found in the phospholipid leaflet can be attributed to difficulties in protein area calculation; however, the corresponding diffusion values, membrane thickness and acyl tail interdigitation in the bilayer provide evidence of a reasonably gel-like bilayer. Furthermore, simulations of an OM transporter, BtuB, were carried out in a symmetric phospholipid bilayer and our model

OM to characterize novel protein-membrane interactions in the native environment.

Significant differences in protein conformation were observed upon comparison of the apo-state BtuB in the symmetric and model OM systems. Reduced fluctuations of the EC loops in the Apo-OM simulations provide strong evidence that LPS plays an important role in OM-protein dynamics. Whether fluctuations would increase on much longer time scales, beyond the 100-200 ns simulated here, is uncertain. Additionally, a clear shift in the arrangement of the EC loops was observed between the symmetric and OM systems. For example, in the symmetric system, EC loop 3/4, which is key to substrate binding, becomes greatly displaced from its substrate-bound conformation due to interactions with phospholipids; in all the OM systems, loop 3/4 maintains an alignment similar to the crystallographic substrate-bound conformation. Surprisingly, the model OM induces an organizational shift in the luminal domain, despite having no direct contact. Lastly, in each of the OM systems, Arg14 becomes unlocked from its interaction with residues in the barrel wall, thus permitting a lower extraction force for removal of the luminal domain as observed from steered MD simulations.

Previous experimental results had determined that calcium binding was required for CBL binding and transport, even though the structural role calcium had on BtuB was not known. The simulations performed here demonstrate that calcium binding promotes a conformational shift in EC loop 3/4 as well as spontaneous secondary structure formation in this loop. Comparing the OM systems reveals that primary loop stabilization occurs by interactions between protein and LPS, with calcium playing a vital role through CBL-binding-site organization and further loop stabilization. We hypothesize that this role provided by calcium ions is necessary for BtuB alone because CBL is the largest substrate among those transported by TonB-dependent transporters.

CHAPTER 3
HYDROGEN MASS REPARTITIONING FOR CHARMM36 MEMBRANE SYSTEMS IN NAMD

Reproduced in part with permission from Curtis Balusek, Hennim Hwang, and James Gumbart. *Hydrogen mass repartitioning for CHARMM36 membrane systems in NAMD*, submitted for publication.

3.1 Introduction

In this chapter we aim to test the application of HMR with a 4-fs time step for membrane systems by comparing lipid properties and lipid-protein interactions. Additionally, the effects of 9-Å cutoff are examined. Several single-lipid, mixed-lipid, and protein-embedded membrane systems were studied. It is found that applying HMR with a 12-Å cutoff provides consistent results in comparison to conventional 2-fs time step and 12-Å cutoff MD across all studied systems. However, employing a 9-Å cutoff altered several structural and kinetic properties for lipid bilayers, even though no significant difference was observed in protein-lipid and protein-protein interactions.

3.2 Methods

Pure Membrane Systems Construction

All-atom lipid bilayers were generated for pure membranes systems using CHARMM-GUI [53]. Three pure-membrane models (2:0, 1:1, 0:2; saturated:unsaturated) of a single lipid-type were generated, as well as a fourth, multiple lipid "Top6" model, with 240 lipids in each leaflet. Each system was solvated then ionized to 150 mM NaCl. Further details about each system and constitutive components can be found in (Table 7.1).

Membrane-Protein Construction

The peptide sequence of the form ac-L_8-nme (L8) was constructed and embedded into the water phase of a box containing a preformed POPC (1-palmitoyl-2-oleoyl-phosphocholine) lipid bilayer with a upper leaflet of 53 lipids and a lower leaflet of 52 lipids. The initial conformation of L8 was an ideal α-helix, placed 10Å from the bilayer surface. The CHARMM36 protein [50] and lipid [45] force fields were used along with TIP3P for water [103] for all equilibrium simulations.

Membrane Mixing

System construction and Anton simulation details of the mixed POPC:cholesterol membrane were described in Hong, *et al.* [104]. The POPC bilayer with 680 lipids were constructed by duplicating four times an equilibrated bilayer with 170 POPC [104]. All simulation conditions of this large POPC bilayer as well as HMR simulations using a 4-fs timestep and 12-Å or 9-Å cutoff of the mixed POPC:cholesterol membrane were identical to those of the pure lipid bilayers described above.

Applied Electric Field Simulations

An OmpF membrane-protein system was created using the crystal structure reported by Yamashita *et al.* [105](PDB: 2ZFG) and embedded in a POPE (1-palmitoyl-2-oleoylphosphatidylethanolamine) bilayer to replicate the systems used in Pezeshki, Chimerel, Bessonov, Winterhalter, and Kleinekathofer. The system contained 99,157 atoms with 176 POPE lipids, 19,421 water molecules, and 420 potassium (K^+) and 396 chloride (Cl^-) ions, to obtain a 1.12 M KCl solution. Three replica simulations of 10 ns each were carried out at 0 V, $\pm$0.2 V, $\pm$0.5 V, and $\pm$1 V mirroring a previous OmpF conductance study [106]. We report the average and standard deviation of the current at each applied voltage by summing up the movement of the charges in the z direction [106, 107, 108]. Results from simulations were limited to 10 ns because a breakdown in the current was observed, especially at

higher fields, such that longer simulations were not utilized.

Molecular Dynamics simulation

After initial construction, HMR was applied to each unique system, thereby creating two identical systems with the only difference being the repartitioned mass. The unmodified system was simulated through conventional MD, with a prototypical 2 fs time step and 12 Å cutoff (2-12) applied with a 10 Å switching function. The HMR systems were simulated using a 4 fs time step and either a 9 Å (4-9) or 12 Å (4-12) LJ-cutoff. Each system was minimized for 1000 steps before being equilibrated for 10 ns. Production was 100 ns for each system with data acquisition on the last 50 ns of simulation. System visualization and analysis was performed with Visual Molecular Dynamics (VMD) [109].

Analysis

Upon completion of the simulations, trajectories were analyzed to measure structural and kinetic properties such as areas per lipid (APL), membrane thickness (D_{HH}), deuterium order parameters (S_{CD}), electron and mass density profiles, compressibility moduli (K_{A}), diffusion coefficient D, and dihedral trans-gauche transition rates.

In the present work, the $\langle$APL$\rangle$ was computed from the size of the simulation box in the x-y plane and standardized by the number of lipids in each leaflet (240). Since we employ pressure coupling, the simulation box is allowed to fluctuate during the simulation, wherein APL was also used to monitor simulation equilibrium. Furthermore, the K_{A} is a measure of the stiffness of the membrane and it was calculated as,

$$K_{\mathrm{A}} = k_{\mathrm{B}}T * A/{\sigma_A}^2 \tag{3.1}$$

where k_{B} is the Boltzmann constant, T is the temperature, A is the area, and ${\sigma_A}^2$ was the variance of the area during the simulation. Membrane thickness is reported as the average

head-to-head distance as measured in the lipid electron density profiles [110, 111, 112].

Deuterium order-parameters (S_{CD}) are used to compare lipid simulations to experimental results of membrane-systems for each acyl-carbon in the aliphatic tail from:

$$S_{CD} =< \frac{1}{2}(3\cos^2\theta - 1) > \tag{3.2}$$

Withing eqn. 3.2, θ is the angle measured between the carbon-hydrogen bond vector and the membrane normal. The bilayer normal vector used for S_{CD} calculation is parallel to the z-axis, because membrane curvature undulations are small in the pure membrane simulations.

Diffusion values are measured from the the mean squared displacement of lipid center of mass motion in the xy-plane.

$$< \Delta r(\tau)^2 >= \sum_{i}^{n} < \|\vec{r}_i(t+\tau) - \vec{r}_i(t)\|^2 >= 2nD\Delta t \tag{3.3}$$

The sum occurs over all molecule mass centers and averaged for all time separations τ. Due to an equilibration phase, the first 1 ns of time lag is discarded. The diffusion coefficient then is proportional to the slope of the linear regime of mean-squared displacement versus time-separation. All diffusion values are measured over the last 50 ns of simulation up to a τ of 20 ns.

The insertion propensity, p_{TM}, of the L8 peptide was calculated as the probability of the peptide being in the TM state. To distinguish the TM state from the surface-associated (S) state, a criterion of $z < 8$ Å was found to be optimal. The free energy of $\mathrm{S} \rightarrow \mathrm{TM}$ partitioning was then calculated as

$$\Delta G_{\mathrm{S}\rightarrow\mathrm{TM}} = +k_{\mathrm{B}}T\log(1/p_{\mathrm{TM}} - 1). \tag{3.4}$$

Sufficient transition events were captured by using elevated temperatures

Radial pair distribution function and clustering analysis of the mixed POPC:cholesterol membrane were performed following Hong, *et al.* [104]. Undulation analysis was performed using the MDAnalysis package [113]. Error analysis was performed following Grossfield, *et al.* [114]. Briefly, a simulation trajectory was divided into M blocks, each of length τ_b. The average of $u^2(q)$ from each block was determined and then used to compute a standard deviation σ_{τ_b}, based on which we obtained the the blocked standard error (BSE):

$$BSE = \frac{\sigma_{\tau_b}}{\sqrt{M}}. \tag{3.5}$$

The error in k_c was then determined by assuming a 'worst-case-scenario combination' of errors from the four wave-numbers analyzed here: we subtracted the BSE from $\langle u^2(q) \rangle$ for the lowest wave-number and added the corresponding the BSEs to $\langle u^2(q) \rangle$ for the remaining three wave-numbers, followed by re-fitting of a first-order polynomial, which produced the upper-bound error in k_c. Conversely, a lower-bound error was obtained. We note that the thus obtained errors were asymmetric, as illustrated in Fig. 7.3.

3.3 Results and Discussion

To determine the effects of HMR, time step, and cutoff on structural and dynamical properties, we performed all-atom MD simulations employing three different simulation schema. As a reference, we first simulated each system studied using a 2-fs time step and 12-Å cutoff with force-based switching from 10-12 Å (referred to as 2-12 throughout the text). The other two schema implement HMR along with a 4-fs time step and either a conventional 12-Å cutoff (referred to as 4-12) or a truncated 9-Åcutoff with no switching (referred to as 4-9). We then report and compare various structural and dynamical properties between each schema.

3.3.1 Pure Membranes

DPPC

Dipalmitoylphosphatidylcholine (DPPC) is a common lipid that has been widely utilized in both experiments and simulation, and thus it has ample data for comparison [115, 116, 45, 117]. Therefore, we used it here as a first test case for determining the possible role of HMR and cutoff on structural and dynamical properties of a pure DPPC membrane. We examined multiple static and kinetic properties of this membrane based on 100 ns simulations, including electron density, area per lipid (APL), area expansion modulus (K_a), order parameters (S_{CD}), diffusion (D), and dihedral trans-gauche transition rates.

First, static properties of the pure DPPC membrane were measured to compare with previous experimental and MD results, beginning with the area per lipid (APL) in each system. APL is utilized to ensure appropriate lipid packing ratios as well as to provide a straightforward comparison to experiment. The APL for DPPC at 323 K in the 2-12, 2-12-HMR, and 4-12 simulations is around 60-61 Å^2 (Table 3.1). Experimental values are slightly larger at 63.3-64.3 Å^2 [117, 5, 115]. As the long-range interactions are reduced in the 4-9 simulation, the area per lipid increases substantially, to 65.5 Å^2. This increase in APL is due to a reduction in the LJ interactions, which are dominant in the aliphatic tails region (Fig. 3.2), also reflected in the decreased order (Fig. 3.1C) and thinning of the electron density profile (Fig. 3.1B). Similar to APL, K_a are measured for each of the systems to determine the membrane area fluctuations. K_a values from each of the DPPC systems (Table 3.1) are in reasonable agreement with the experimentally observed value (231 mN/m) with the 2-12-HMR system showing the lowest percent difference ($\sim 1.5\%$).

Next we look at order parameters (Fig. 3.1C), and the values from our 2-12 simulation agree well with previous experimental [115] and simulation [118, 119, 120, 45] results. When comparing the 4-12 simulation to the 2-12 simulation, there is a small decrease in lipid order resulting in a $5.7 \pm 2.8\%$ difference on average over all the carbon positions;

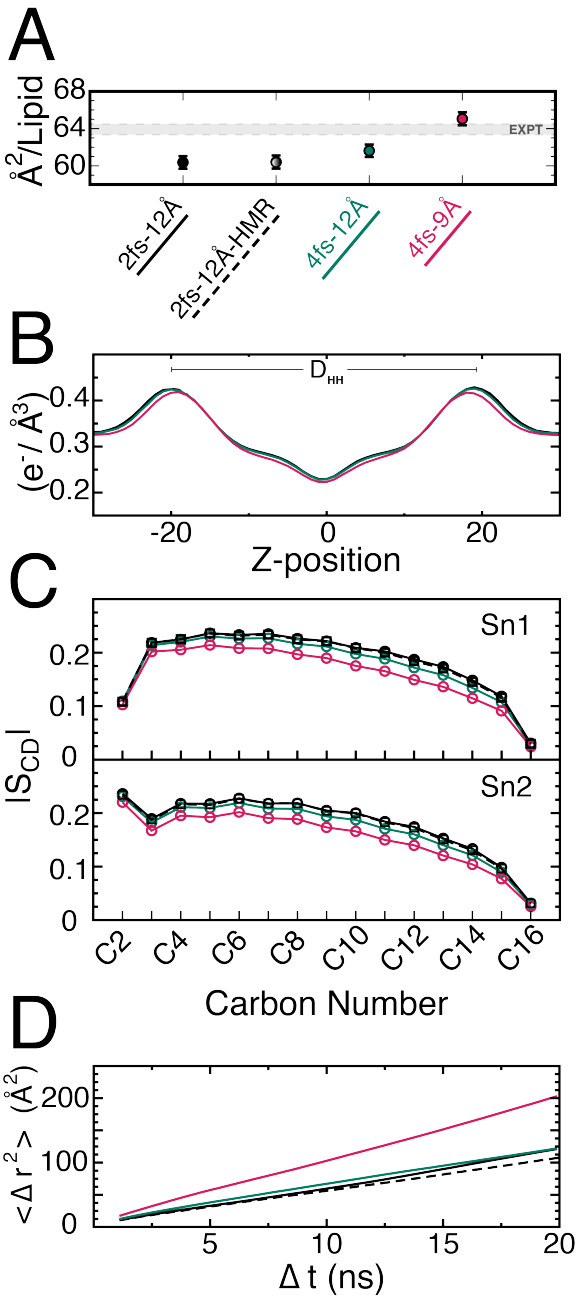


Figure 3.1: DPPC membrane properties. For each panel, 2-12 is colored in black, 4-12 in green, and 4-9 in red. An additional 2-12 with HMR is also shown as a dashed black line. A. Area per lipid with standard deviation bars taken from last 50 ns of production. B. Plot of electron density; membrane thickness (D_{HH}) is measured from left peak to right peak of each distribution. C. Plot of sn1 (top) and sn2 (bottom) lipid order parameters (circles) for the 2-12 (black), 4-12 (green), and 4-9 (red) simulations. The 2-12-HMR values are shown as black squares. D. Mean-squared displacement vs. time averaged over all lipids and times for each simulation.

when comparing 2-12 and 4-9 simulations, the decrease ($17.4 \pm 6.4\%$) is much larger, suggesting that the cutoff has a significant effect on lipid order. To discriminate between HMR and the 4-fs time step, we performed an additional simulation with 2-fs time step and 12-Å cutoff as well as HMR (2-12-HMR), which showed a decrease of $1.0 \pm 0.8\%$ on average in the order parameters (Fig. 3.1C), similar to the 2-12 simulation. This result suggests that HMR causes little to no effect in the lipid order parameters, an increased time step causes a slight but measurable decrease in order, and a reduced cutoff causes a substantial decrease in order. The decrease in order is also evident from an increased interdigitation of the aliphatic tails in 4-9 simulation (Fig. 3.2, bottom) compared to 2-12 (Fig. 3.2, top) and 4-12 (Fig. 3.2, middle) simulations. Furthermore, aliphatic interdigitation between membrane leaflets results in membrane thinning by 1 Å in the 4-9 simulation compared to the 2-12 simulation (Fig. 3.1B).

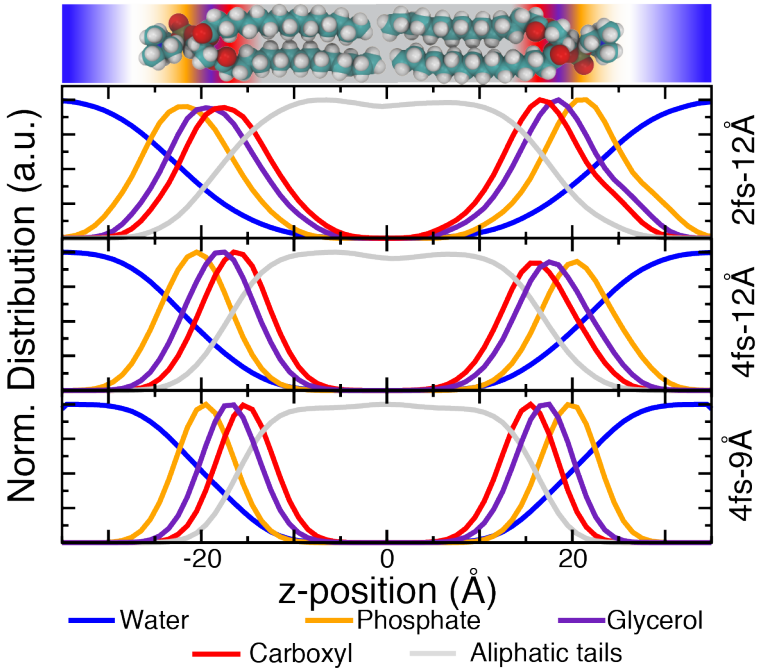


Figure 3.2: Plot of molecular densities of DPPC in the 2-12 (top), 4-12 (middle), and 4-9 (bottom) simulations. The overlap in aliphatic tails (grey) is shown to increase from 2-12 and 4-12 to 4-9 suggesting that the shorter (9Å) cutoff compresses and interdigitates the lipid tails more than the 12Å cutoff.

Next, we examined kinetic properties for each simulation, such as the rate of lipid diffu-

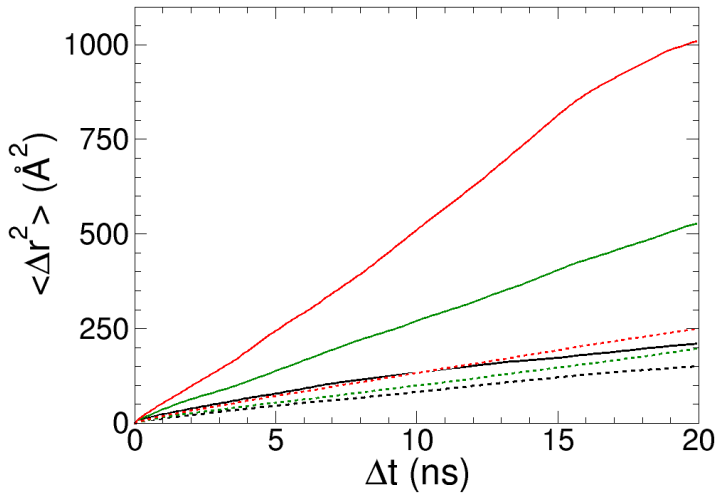


Figure 3.3: Diffusion plot of DPPC simulated using Lowe-Anderson thermostat (solid) with the langevin piston simulations (dashed) as reference. The 2-12 simulations are shown in black, 4-12 in green and 4-9 in red. Applying the Lowe-Anderson thermostat results in significantly larger diffusion values across all simulation protocols. Increases in diffusion are approximately 70% in for the 2-12 system and grow to 300% when comparing the 4-9 system.

sion and dihedral transition rates for the aliphatic tails. While diffusion values from simulation are not reliable for a number of reasons, including a dependency on box size [121] and thermostat [32], they provide a useful comparison between simulation schema. Diffusion values (D) are calculated from the slope of the mean-square displacement vs. time, averaged over lipids and time (Fig. 3.1D). For simulations with a 12-Å cutoff (2-12, 2-12-HMR, 4-12), diffusion values ranged from 1.30-1.64 Å^2/ns, slightly smaller than the experimental value of 1.78 Å^2/ns [122, 123]. For the 4-9 simulation, D was much larger at 2.53 Å^2/ns. However, it has been demonstrated that Langevin dynamics, used in all simulations here to control the temperature, reduces D by 35% [32]. Therefore, we also ran the 2-12 simulation using the Lowe-Andersen thermostat instead, yet found that D was significantly larger at 2.47 Å^2/ns (Fig. 3.3). The increase in D was even greater for 4-12 at 2.47 Å^2/ns and 4-9 at 6.56 Å^2/ns.

To determine the source of the dihedral transition rate of aliphatic tails were measured

in 2-12 and 2-12-HMR. The trans-gauche (t-g) transition rate for the aliphatic dihedral angles were measured. It has been postulated that intramolecular conformation of the lipid molecule is closely related to the intermolecular structure of the membrane, such that the t-g transition in dihedral torsional may be a good measure for equilibration time needed [124]. This was done by measuring the rate of transition between the trans, gauche+ and gauche- conformations of sequential carbons on the lipid tails of DPPC. It was found that even when measuring the transition rates between trans, gauche+, and gauche- were nearly identical across all simulation schema over a 1 ns simulation performed with 10 fs frame output. More specifically, the average values for each dihedral angle was within $1\,\%$ of the other simulation protocols.

In summary for the test case of DPPC, simulations suggest that HMR has a negligible effect on the structural and kinetic properties when comparing the 2-12 and 2-12-HMR systems. Second, the longer time step has a slight effect on the membrane, in particular, the $\sim 5\%$ difference order parameters comparing 2-12 and 4-12 indicate a more fluid bilayer. Lastly, as is evident in each measurement from 4-9, the shorter cutoffs lead to more fluid lipids ($\sim 17\%$ difference from 2-12), membrane thinning ($\sim$ 1Å shorter than 2-12), more diffusive lipids ($\sim 40\%$ increase from 2-12 and 4-12), and a change in the overall lipid density.

Other Membranes

Results of lipid membrane simulations are usually compared with observations from X-ray or neutron scattering, or NMR experiments by considering temporal and spatial averages of various observables, e.g., bilayer thickness (D_{HH}) or APL [5, 131]. After measuring the effects of HMR, longer time-step, and a shorter cutoff on an unsaturated membrane, we proceeded to simulate three additional membranes to determine if varing lipid composition would alter our initial observations. We performed simulations on 1-palmitoyl,2-oleoyl-sn-glycero-phosphatidylcholine (POPC, one unsaturated tail), 1,2-oleoyl-snglycero-

System	Setting	APL (Å^2)	D_{HH} (Å)	K_{A} (mN/m)	D (Å^2/ns)
DPPC (323 K)	2fs-12Å	60.37 $\pm$0.67	39.05	249.2	1.45
	2fs-12Å HMR	60.41 $\pm$0.70	38.96	227.7	1.27
	4fs-12Å	61.62 $\pm$0.67	38.73	252.5	1.43
	4fs-9Å	65.1 $\pm$0.69	38.04	250.9	2.42
	Exp.	63.3 – 64.3[115, 116]	37.8[116]	231[5] (318K)	1.78[122]
POPE (310 K)	2fs-12Å	57.36 $\pm$0.59	40.4	281.9	0.64
	4fs-12Å	57.96 $\pm$0.67	40.3	222.4	0.97
	4fs-9Å	61.07 $\pm$0.61	39.5	279.7	1.18
	Exp.	59.8 – 60.8[125]	40.0[125] (303K)	233[126]	–
DOPC (295 K)	2fs-12Å	66.95 $\pm$0.57	38.7	350.3	0.65
	4fs-12Å	66.97 $\pm$0.62	38.7	288.6	0.61
	4fs-9Å	70.31 $\pm$0.71	38.0	241.4	1.13
	Exp.	67.4[127] (303K)	36.7[128] (303K)	265[129] (303K)	1.0[130] (296.5K)
Top6 (310 K)	2fs-12Å	61.54 $\pm$0.57	37.2	327.4	0.78
	4fs-12Å	62.46 $\pm$0.63	37.0	269.0	0.87
	4fs-9Å	64.75 $\pm$0.62	36.9	290.0	1.41
	Exp.	63.0 $\pm$0.19[74]	–	240[74] (310K)	–

Table 3.1: Average properties of lipid bilayer; Area per lipid (APL), membrane thickness (D_{HH}), area compressibility modulus (K_{A}), and diffusion coefficient (D). The values of temperature in the parentheses indicate that the property was measured at that specific temperature.

phosphotidylcholine (DOPC, two unsaturated tails), as well as the so-called "Top6" membrane, which is a mixture of saturated, unsaturated, and cyclic-containing lipids. In order to expand upon the same measurements as the DPPC simulations, each of these membranes were simulated using the 2-12 schema as a control, as well as the additional 4-12 and 4-9 simulations applying HMR. We report the various physical properties of the membranes in Table. 3.1. These properties were calculated over the last 50 ns of each 100 ns simulation.

For each system, APL values in the 2-12 and 4-12 systems are in a good agreement with experimental measurements (See Fig. 7.1 for graphs of APL fluctuation over the entire simulation period). It was also observed that the 4-9 systems have the largest APL values in each of the systems, indicating that APL increases with a shorter cutoff (9 Å vs. 12 Å). Furthermore, because the membrane thickness does not vary considerably between each of the schemas, there is a resulting increase in membrane volume for each of the 4-9 systems.

One particular result of interest is the difference in diffusion values observed in each of the pure membrane systems. There is a small increase going from 2 fs to 4 fs; however, each 4-9 system has a diffusion value approximately 40% greater than in either the 2-12 or 4-12. This is due to the increased fluidity of the membrane as is measured from the lipid order parameters (see SI). However, unlike all other properties, the values of K_{A} showed no consistency between simulation schema. It is known that K_{A} values from MD simulation can be sensitive to the time frame used for measurement, because it is calculated based on the area fluctuation, the discrepancy with experimental values were expected.

Lipid Mixing and Membrane Bending Modulus

To examine the impact of time step, HMR and cutoff on lipid mixing, we turn to the POPC:cholesterol mixture previously investigated by microsecond Anton simulation[104]. Each leaflet of the mixture was composed of 70 POPC and 35 cholesterol, with the latter lipid species initially placed at the center of the bilayer. This mixed membrane was simulated for 1 μs with either a 4-9 or a 4-12 time step-cutoff combination. The Anton

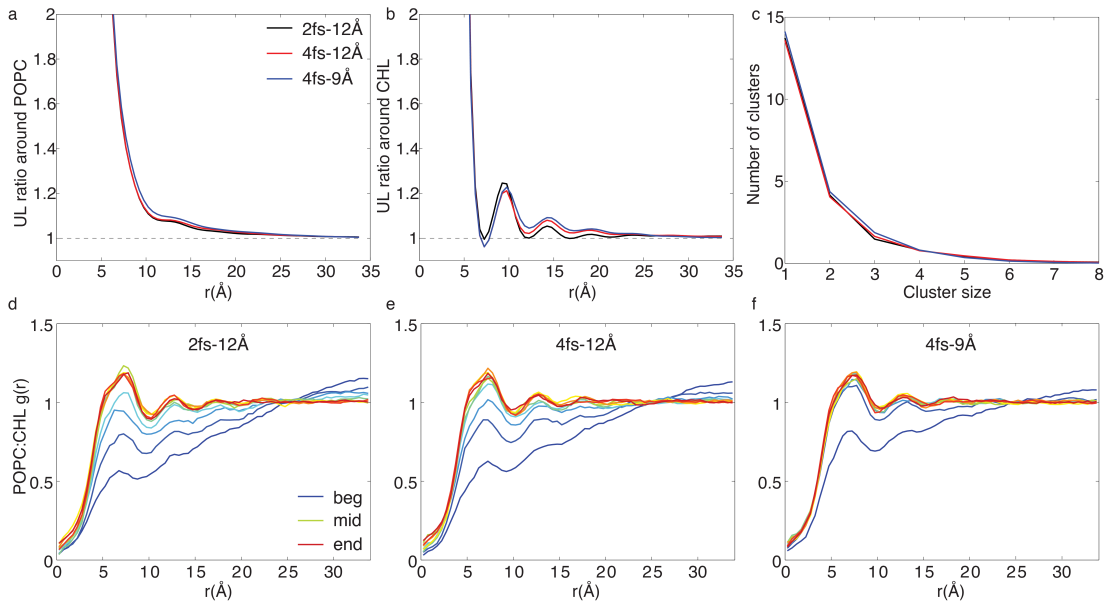


Figure 3.4: Impact of HMR, time step and cutoff on lipid mixing. (a-b) Average ratio of unlike neighbors to like neighbors (UL) around a given lipid species. The expected values based on mixing ratio are indicated by dashed lines. (c) Clustering analysis results of the POPC:CHL mixture. (d-f) Time evolution of the radial pair distribution functions g(r) for POPC:CHL. g(r) is averaged in 100-ns blocks and colored by simulation time, with blue, green and red indicating the beginning, the middle and the end of a simulation, respectively.

simulation reported previously [104] provides the reference for a 2-12 time step-cutoff combination. Unless otherwise noted, we analyzed the trajectory from the first microsecond of the 2-μs Anton simulation for a fair comparison with the 1-μs runs performed in this work. As shown in Fig 3.4, regardless of the time step-cutoff combinations, the final structures of the mixtures are similar to each other, as reflected by the ratios of unlike-to-like (UL) neighbors (Fig 3.4 a,b) and the size distribution of cholesterol clusters (Fig 3.4 c). Therefore, the equilibrium distribution of the lipids is unaffected by the choice of time step, use of HMR, or cutoff. The APL of the mixtures, however, is clearly affected: averaging over all POPC and cholesterol yields an APL of 46.4 Å^2, 47.1 Å^2 and 49.4 Å^2 for the 2-12, 4-12, and 4-9 simulations, respectively, reflecting a trend consistent to that observed in our pure lipid simulations (Table 3.1). Comparison of the three simulations also reveals a clear difference in the speed of mixing. Semi-quantitatively, this can be seen from Fig 3.4 d-f and Fig 7.2, which show the evolution of the radial pair distribution func-

tion g(r) over the course of the 4-9 and the 4-12 simulations as well as the first 1-μs of the previously performed 2-12 simulation. The g(r) curves, drawn every 100 ns, suggest that the 4-9 run converges much faster than the other two simulations. Quantitatively, the lateral diffusion coefficient of cholesterol in the 4-9 simulation (6.5×10^{-8} cm^2/s) is over two times faster than that in the 2-12 simulation (2.5×10^{-8} cm^2/s). Similarly, POPC diffuses significantly faster in the former system, with a diffusion coefficient of 6.1×10^{-8} cm^2/s (2.2×10^{-8} cm^2/s in the 2-12 run). In the 4-12 simulation, diffusion of cholesterol (4.4×10^{-8} cm^2/s) and POPC (3.0×10^{-8} cm^2/s) is also accelerated compared with the 2-12 simulation, although the acceleration is much smaller than in the 4-9 simulation, thereby, resulting in distinct g(r) evolution profiles between the two 4-fs simulations. Apart from lipid lateral diffusion, it is worth noting that a cholesterol flip-flop event was recorded in both the 4-9 and the 4-12 simulation, in contrast to zero flip-flop event recorded throughout the 2-μs 2-12 simulation. The larger APL in the 4-fs simulations may have contributed to their increased cholesterol flip-flop events, as the former is linked to the free energy barrier of defect formation in a membrane [132]. Overall, our data indicates that the choice of time step, cutoff and use of HMR has a negligible effect on the distribution of lipids in an equilibrated mixture, although other equilibrium properties, such as the APL, are evidently affected. On the speed of lipid mixing, while increasing the simulation time step from 2 fs to 4 fs and applying HMR has a rather moderate effect, decreasing the cutoff from 12 Å to 9 Å can significantly accelerate mixing.

Following our investigation on lipid mixing, we went on to evaluate how HMR may affect one of the most important material moduli of a membrane, namely, its bending modulus (k_c). Using a bilayer with 680 POPC, we performed three 1-μs simulations with 2-12, 4-12, and 4-9 time step-cutoff combinations, respectively. The $\sim$150 Å$\times$150 Å bilayer supported relatively long-wavelength undulation modes, thereby, allowing us to determine k_c from $\langle u^2(q)\rangle$, the average square amplitude of undulation at a given wavenumber q. More specifically, k_c was calculated from the last 900 ns of the 1-μ trajectories according

to $\langle u^2(q)\rangle = k_c k_B T A^{-1} q^{-4}$ using the MDAnalysis package [113] and methods presented in Refs [133, 134]. As shown in Fig 7.3, k_c was found to be 30.9 $k_B T$ (12.9$\times 10^{-20}$ J) in the 2-12 simulation, which is comparable to the experimental value (9.0$\times 10^{-20}$ J) reported for a similar lipid bilayer (1-stearoyl-2-oleoyl-*sn*-glycero-3-phosphocholine) [135]. To estimate the uncertainty in k_c, we first examined the blocked standard error in $\langle u^2(q)\rangle$. The undulation mode with the smallest wavenumber (longest wavelength) was found to have the largest error, and this error was approximately an order of magnitude greater than that shown in Fig 7.3, if, instead of the last 900 ns, only the last 50 ns of trajectories were used in the analysis (data not shown). This behavior supports the need for microsecond-long trajectories in reliable analysis of k_c. Here, the uncertainty in our k_c values was found to be approximately 2 k_BT. Compared with the 2-12 run, k_c decreased slightly to 28.6 $k_B T$ in the 4-12 simulation. In the 4-9 simulation, a further decrease was seen, with k_c reaching 25.4 $k_B T$. We note that the difference between the 2-12 and the 4-9 simulations is well beyond the estimated uncertainty in k_c, indicating that the comparison is statistically meaningful. The average projected APL was found to be 64.2, 64.9 and 67.3 Å^2 in the 2-12, 4-12, and 4-9 simulations, respectively. Taken together, these results again reflected the weakened lipid interactions when a short cutoff (9 Å) was adopted. Such weakened interactions not only produced an increased area per lipid, but also reduced the energetic cost of bending the membrane, thereby, resulting in a decreased k_c.

3.3.2 Membrane-Protein Systems

Electric Field Simulations of OmpF

One of the most important functions of membrane proteins is the channeling of ions into and out of the cell. OmpF is a well-characterized trimeric protein that acts as a nonspecific ion channel in the outer membrane of Gram-negative bacteria [136, 137]. To continue the investigation into the effect of timestep, cutoff, and HMR on protein properties like conductance, we performed simulations with an applied electric field on the outer membrane

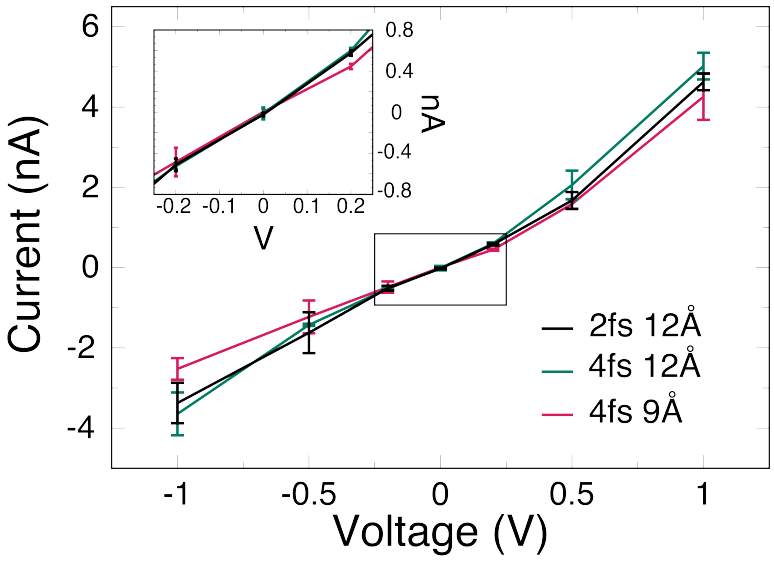


Figure 3.5: Current vs. voltage for applied field simulations of OmpF in POPE membrane. 2-12 is shown in black, 4-12 in green, and 4-9 in red. The inset graph magnifies results from lower potentials. Differences in timestep or cutoff and little distinguishable difference in measured current in each system.

ion channel OmpF. Scaling by molarity, our results across all simulation schema compare favorably to those reported by Pezeshki et al. [106] including the higher current for positive voltages due to a slight cation selectivity of the channel. At low voltages we observe a near exact agreement between all three simulation protocols (Fig. 3.5). At higher voltages, 4-9 underestimates the current, while 4-12 slightly overestimates the current with respect to the conventional 2-12. Furthermore, the deviation observed in the 4-9 system is higher than that observed by 4-12 compared to 2-12 at higher applied potentials. However, the deviation for each of the HMR systems is within, or very nearly within the error of the conventional 2-12 system indicating that the effect of HMR, longer timestep, or shorter cutoffs on the membrane dynamics does not significantly effect the dynamics of OmpF and therefore the conductance is not significantly influenced.

Octoleucine Insertion in POPC

To further investigate how membrane-protein kinetics are affected by HMR, we performed long equilibrium simulations of an octoleucine (L8) helix embedded within a symmet-

ric POPC bilayer (see Methods) using a 4-fs timestep, HMR, and both 12-Å and 9-Å Lennard-Jones cutoffs. With the peptide starting in a transmembrane (TM) state and using a high temperature (423 K), we measured the number of transitions between the TM and surface-associated (S) states over the course of a 1.3- and 1.5-μs production run for 12-Å and 9-Å cutoffs, respectively. The TM state was defined as $|z| < 8$ Å, where z is the distance between the center of mass of the peptide backbone and the center of the membrane, consistent with previous studies of L8 insertion into a lipid bilayer [138]. The high temperature was necessary for sufficient sampling of the two states, and backbone dihedral restraints were added to ensure the peptide did not unfold at this temperature [138]. We note that the 200 kcal/mol dihedral force constant used in the non-HMR systems produced instabilities in the HMR system due to the smaller masses of the heavy backbone atoms compared to the non-HMR system. Backbone N and C_α masses are reduced by $\sim$15% due to HMR, therefore in order to maintain roughly the same mass-to-force-constant ratio, we similarly reduced the dihedral force constant to 175 kcal/mol in the HMR system. We observed 14 transitions in 1.3 μs (10.7 transitions/μs) with a 12-Å cutoff, while we only observed 10 transitions in 1.5 μs (6.7 transitions/μs) with a 9-Å cutoff (see Fig. 3.6A). In addition, the peptide spent far less time in the S state for the latter, with $\Delta G_{\mathrm{S}\to\mathrm{TM}} = -2.8$ kcal/mol for the 9-Å cutoff compared to -0.9 kcal/mol for the the 12-Å cutoff (see Fig. 3.6B and Methods for calculation $\Delta G_{\mathrm{S}\to\mathrm{TM}}$). Previous multi-$\mu$s-scale simulations of L8 at 423 K run on the Anton supercomputer with a 2.5-fs timestep and a 13–14-Å cutoff without HMR produced roughly 12 transitions/μs, with a temperature-independent $\Delta G_{\mathrm{S}\to\mathrm{TM}} = -0.9 \pm 0.2$ kcal/mol [138]. HMR systems with a 12-Å cutoff are in good agreement with these results, reproducing the free energy difference and only slightly underestimating the S→TM transition rate. HMR with a 9-Å cutoff, however, significantly underestimates both the free energy difference and the transition rate.

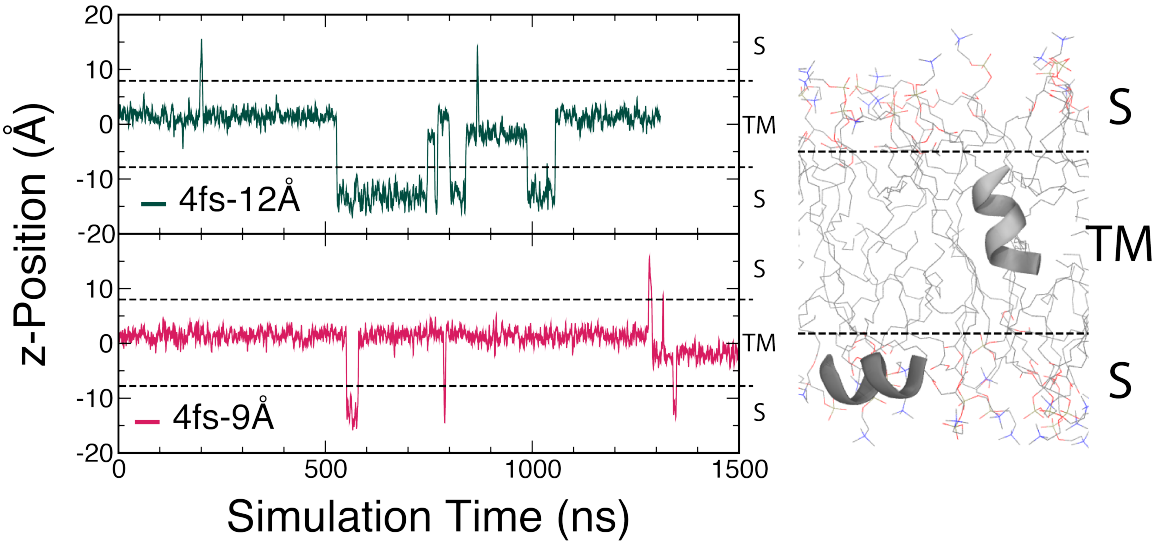


Figure 3.6: Transmembrane to surface-associated transitions of L8 helix in a POPC bilayer with HMR. (Left) Position of helix within the membrane using HMR with a 4-fs timestep and either a (top) 12-Å or (bottom) 9-Å Lennard-Jones cutoff. Center of the membrane is defined as $z = 0$ Å. (Right) Position and orientations of L8 in surface-associated (S) and transmembrane (TM) states. Peptide shown in cartoon representation and colored blue. Lipid molecules are shown in line representation and colored by atom name (hydrogen atoms ommitted). TM states are defined as $|z| < 8$ Å.

Benchmarks

Benchmarking simulations were carried out for the DPPC membrane (113,064 atoms) and the large, 680-lipid POPC membrane (170,844 atoms). Each system was run on 1, 2, 4, 8, 12, 16, and, for the larger system, 24 CPU-only nodes of (1) Stampede2 at Texas Advanced Computing Center (Intel Xeon Skylake CPUs; 48 cores/node) and (2) Bridges at Pittsburgh Supercomputing Center (Intel Haswell CPUs; 28 cores/node). Each simulation system was run using (1) a 2-fs time step and 12-Å cutoff, (2) a 4-fs time step and a 12-Å cutoff, (3) a 4-fs time step, a 12-Å cutoff, and PME evaluated every 8 fs, (4) 4-fs time step and a 9-Å cutoff, and (5) 4-fs time step, a 9-Å cutoff, and PME evaluated every 8 fs.

As expected, the 4-fs/9-Å/8-fs-PME simulations had the greatest performance (measured in ns/day) at practically all node counts. However, the benefits accruing from each approximation were not equal. The 4 fs-12 Å simulations are consistently $\sim$75% faster

than the 2 fs-12 Å simulations on CPUs. However, employeeing a 9-Å cutoff only speeds up simulations by an additional 20-45%. Evaluating PME electrostatics every 8 fs instead of 4 fs is of mixed benefit at either cutoff, giving at most 40% improvement in speed at high node counts; at reasonable (efficiency $> 75\%$) node counts, the speedup is 20% at most.

Although no production simulations in this paper used GPUs, we also benchmarked on 1-4 NVIDIA GTX 980 cards on a single node with two Intel Xeon Haswell CPUs (24 cores, using 6 cores/GPU). Going from 2-fs to 4-fs time steps gives a consistent speedup of 40%. No improvement was seen when shortening the cutoff; short-range interactions are evaluated on the GPU(s), but these simulations are CPU-limited. Additionally, no benefit was seen when evaluating PME every 8 fs instead of 4 fs.

3.4 Conclusion

All-atom MD simulations of lipid-containing systems are complicated and computationally burdensome due to the typically large ($> 100k$ atom) systems necessary for statistical convergence of data. As such, conventional equilibrium MD simulations have utilized a number of methods to speed up computational dynamics at the cost of resolution, e.g. united atom force fields or coarse-grain, or through the introduction of restraints, i.e. rigid bonds. Here, a method known as hydrogen mass repartitioning (HMR), is studied as applied to lipids. This method has previously been validated for simulating proteins but not for lipids. Here, we validate the application of HMR in comparison to conventional MD as well as with a larger integration time step and shorter long-range cutoffs. From simulation data, it is shown that HMR has a negligible effect under conventional MD. Simulations performed with a 2 fs time step and 12 Å cutoff with and without HMR produce almost identical structural and kinetic properties of a pure lipid membrane. Similarly, HMR with a longer time step has a slight difference in membrane order parameters, resulting in a slightly more fluid membrane, and diffusion values are approximately the same or greater by $< 5\%$. However, utilization of shorter cutoffs create much more fluidized lipid membranes, as seen

in lipid densities, membrane thickness, and order parameters, as well as faster diffusing lipids. In the context of pure membranes, the utilization of a short, 9 Å cutoff is not advised; however, simulations with associated or embedded proteins suggest that the shorter cutoff does not create substantial differences in production values.

CHAPTER 4
A METHOD FOR CONSTRUCTING A CELL-ENVELOPE MODEL OF GRAM-NEGATIVE BACTERIA

4.1 Motivation

The cell envelope of Gram-negative bacteria is complex in its composition of two membranes, a cell wall, and constituent proteins. As described in previous chapters, the asymmetric OM is primarily composed of $\beta-$barrel proteins whereas the symmetric IM contains $\alpha-$helical proteins. Secondly the Gram-negative cell wall is much thinner than the Gram-positive counterpart and is linked to the OM by membrane associated lipoproteins. Contemporary work performed on the OM and attached cell wall have elucidated stress-strain relationships between the two through attachment of Braun's Lipoprotein [139]. From chapter 3, we see that simulation speed increases can be achieved utilizing hydrogen mass repartitioning, for longer simulations in the same amount of computational time. Lastly, there have been innumerable studies published describing the dynamics of and membrane effects of IM proteins.

Until recently, protein systems studied using MD have typically been performed in isolated environments.This systematic reduction for studying proteins is two-fold. The first reduction has largely been due to limitations in computational resources, necessitating small atom counts in order to get long simulation-time dynamics; however, with the ongoing improvement of super computing resources, such as Anton, Blue Waters, Stampede, etc., large ($> 10\,M$ atom) systems are becoming an increasing reality [140, 141, 142]. Second, it is often rigorous to elucidate predictive information from a single-protein containing simulation let alone a multi-protein system; however, great strides have recently been published to describe the behavior of protein complexes [76, 143, 144, 142]. Another consistent issue is

the amount of available experimental information for various Gram-negative species; from available crystal structures, protein localizations, relative abundances, and intracellular interactions. This is particularly evident for species other than *E. coli*.

Pursuant to the ongoing development and study of protein-complexes, the natural question to ask is if there are protein complexes that span the cellular envelope of Gram-negative bacteria. In short, yes, there are multiple protein complexes of distinct function and interest, which contribute to various processes from protein and lipid insertion, e.g. $\beta-$barrel assembly machinery and lipopolysaccharide insertion machinery respectively, to secretion systems, i.e. AcrA/B-TolC multi-drug efflux pump (see Fig. 4.1) to name a select few. As such, it is import to construct an all-atom model which contains all pertinent interactions for the multi-protein complex being considered. Therefore, we describe a method for utilizing genomic and proteomic data for identifying, localizing, and quantifying proteins by relative abundance as well as using experimental mass spectrometry data to co-localize periplasm proteins by interaction. This method provides a means to construct reliable all-atom MD systems with relative abundances specifically related to envelope-spanning complexes. This method is validated for K12 *E. coli* for a generic periplasm.

4.2 Method

First, whole-cell genomic data is obtained from Uniprot [145, 146] to establish gene identification as well as a thorough list of protein localizations. Since the protein localizations are incomplete in Uniprot, an *E. coli* K12 specific database is used to validate and reduce the initial genomic dataset from Uniprot [147]. An external database may not be available for other Gram-negative species and special interest should be considered when highly abundant proteins exist with no structure. Next, whole-cell relative abundances are obtained from experimental proteomics studies [148]. For *E. coli* K12, there are currently thirteen unique whole-cell protein abundance surveys. Of those thirteen, four have a whole-cell coverage $> 90\%$, including a WholeCell [148] integration of multiple methods, Absolute

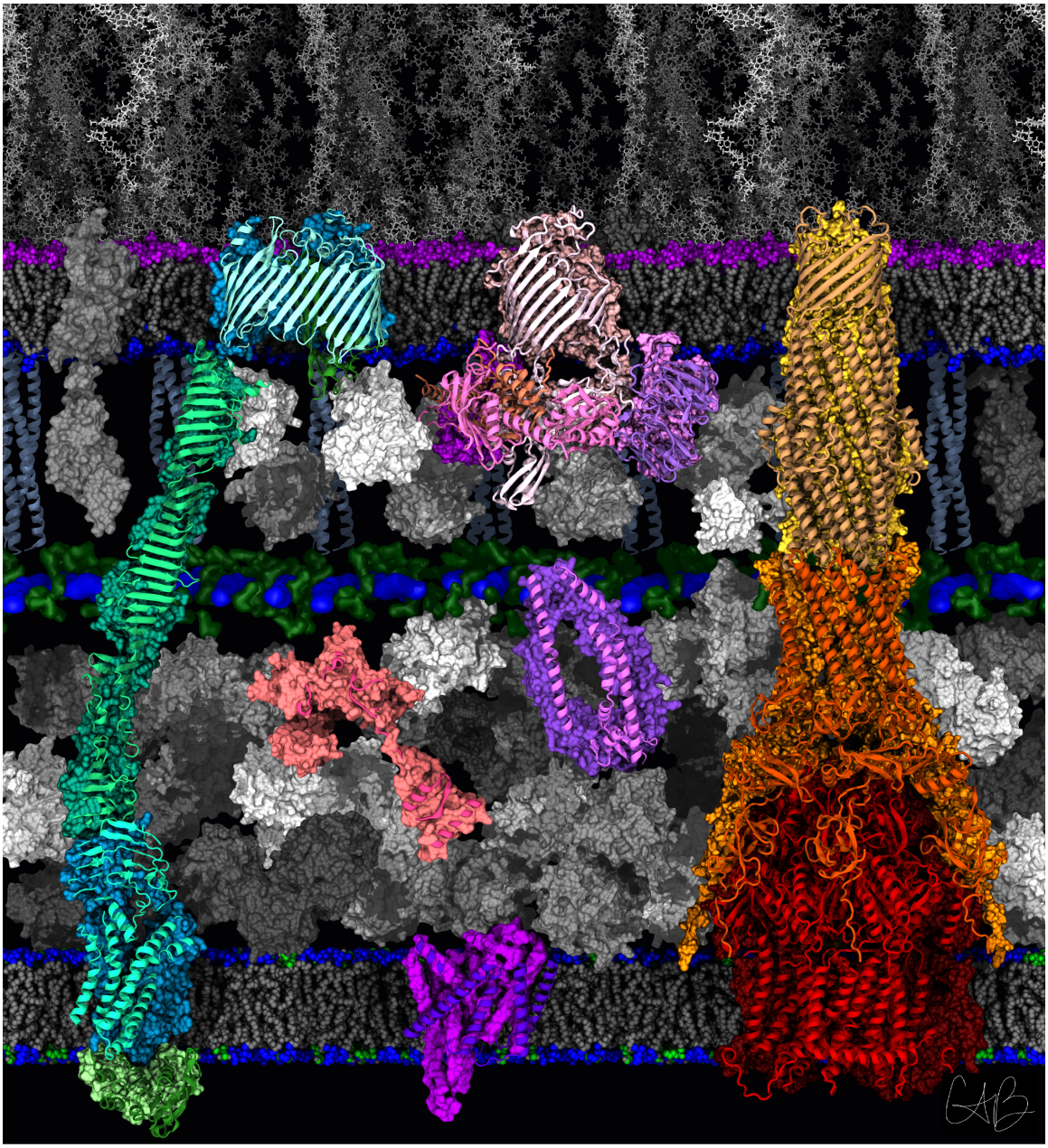


Figure 4.1: An all-atom rendering of the Gram-negative cellular envelope, showcasing various envelope-spanning protein complexes. The extracellular space is along the top and cytoplasm along the bottom. The LPS insertion machinery (Lpt-complex) is shown in various green-blue hues at left. The $\beta-$barrel insertion machinery and associated chaperones are located centrally in purple shades. AcrA/B-TolC efflux pump is shown in Orange-Red-Yellow, respectively, on the right. The cell wall is centrally located in green and blue volumes and connected to the OM by crystal-blue $\alpha-$helix lipoproteins. Periplasm proteins (various grey volumes) fill the space between the IM & OM. Lipid tails are shown as grey spheres and phosphate head-groups are in blue/green spheres for the IM. For the OM, phosphate head groups are blue spheres and magenta spheres designate the amphiphilic boundary of LPS.

protein expression profiling (Apex12) [149], Exponentially modified protein abundance index, (emPAI12) [149], Intensity Based Absolute Quantitation (iBAQ12) [149]) and another three have abundances in excess of 60%, LC-MS/MS (GPM14) [150, 151], LC-MS/MS (Krug13) [152], Krug-emPAI [152]. Each dataset provides a relative abundance of proteins included in the survey normalized by protein mass. For example if protein A has an abundance of 11000 and protein B 1100 within a specific survey, the relative ratio of A:B is 10:1. Special care must be utilized to ensure abundance percentages are reported for comparison between abundance studies. This process provides a quantifiable method to procure proteins by localization in the cell-envelope, e.g. extracellular, OM, IM, periplasm, cell-wall, or membrane associated; as well as by relative abundance.

As an additional means of producing accurate envelope models, a recent experimental method for protein co-localization has been explored by the Freddolino lab at the University of Michigan. This process utilizes cross-linking and liquid chromatography tandem mass spectrometry (LC-MS/MS) to determine protein interaction networks within the cell-envelope. This process is performed by functionalizing exposed amino-acid side-chains of cell envelope proteins with a unique linker then observing the spectroscopic mass shift from cross-linking events. Since proteins have unique masses, observed spectroscopic shifts are a unique linear combination of those masses, and therefore easily identified from proteomic information. Currently, results from their cross-linking measurements are unpublished and are thus not included here.

4.3 Results

In order to construct a cell-envelope model of *E. coli*, the most abundant proteins in the periplasm and OM are utilized. The results from each localization can be seen in Tables 4.1, 4.2, 4.3, 4.4. Selections for the most abundant proteins was selected for average abundance in excess of 100 parts per million or the most abundant 20 proteins in the localization.

For the OM, the most abundant protein is OmpA, an 8-stranded $\beta-$barrel protein, with a periplasm domain thought to associate with the cell wall [153]. This protein has an abundance about four times greater than the next greatest OM proteins, OmpX, OmpF, OmpT, and OmpC, shown in Table 4.1. OmpF and OmpC are proteins that trimerize and permit passive diffusion of ions across the OM [136, 137, 154], such that an accurate model with a trimeric state of each of these proteins should include four OmpA each, three OmpX, and three OmpT. Furthermore, it is believed that the OM is approximately $\sim 50\%$ protein by mass [155], such that the lipid composition can be postulated. The approximate mass (Daltons) of K12 LPS is 3.9 kDa, and a typical phospholipid (POPE) is 0.720 kDa. There is a 3:1 ratio of phospholipids to LPS due to size, such that a unit of lipids is around 6 kDa. Each of the OM associated proteins vary in size and mass; however, utilizing the most abundant, OmpA is 35 kDa, OmpX is 16 kDa, OmpF is 37 kDa, OmpC is 38 kDa, and OmpT is 33 kDa. An outer membrane model put forth consists of proportional parts OmpA, OmpX, OmpF, OmpC, and OmpT that has a ratio of 12:3:1 trimer:1 trimer:3 $\beta-$barrels based upon abundance and a corresponding 132 lipid units (132 LPS and 396 phospholipids). Similarly, OM-associated lipoproteins, seen in Table 4.2, have abundances that are approximately similar to that of the OM $\beta-$barrel proteins. The most abundant is Braun's lipprotein, lpp, which is known to be lipidated at the C-terminus, and inserted into the inner leaflet of the OM. Furthermore, the trimeric state of lpp has a monomeric glycosidic cross-link with the peptidoglycan cell-wall [72, 73, 74]. The next two most abundant lipoproteins, ecnB and osmE are associated with regulation of osmotic shock of the bacterium, but unfortunately have no crystal structure for *E. coli* K12.

The next area of greatest interest lies in constructing an accurate periplasm model. Due to the substantial barrier to entry afforded by the OM, imagining the constitutive components of the periplasm is virtually impossible. This is especially difficult whe attempting to determine the amount of protein encapsulated by the periplasm. A recent paper by Sochacki *et al.* made a strong argument for estimating the periplasm content by subjecting

Locus	Gene	WholeCell	Apex12	emPAI12	iBAQ12	GPM14	Krug13	Krug-emPAI	Avg (ppm)
OM beta-barrel proteins (13 of 65)									
b0957	ompA	2767	1452	1080	1435	6337	14872	150	4013
b0814	ompX	546	0	0	0	2463	4230	128	1052
b0929	ompF	695	1329	671	856	2380	1341	11	1040
b0565	ompT	560	290	212	174	1389	3698	266	941
b2215	ompC	679	0	354	413	2590	2489	28	936
b2000	flu	299	995	475	699	227	285	3	426
b0584	fepA	168	0	0	0	559	1604	24	336
b3035	tolC	219	60	133	70	586	1126	19	316
b0177	bamA	184	221	132	167	384	577	12	240
b2155	cirA	259	0	0	0	363	965	31	231
b4036	lamB	102	300	65	96	389	4	0	137
b0411	tsx	125	66	78	94	357	235	3	137
b1782	mipA	86	0	94	62	307	213	9	110

Table 4.1: Most abundant integral outer membrane proteins (abundances $>$ 100 ppm). There are a total of 65 known proteins to be integral outer membrane proteins in *E. coli*. The most abundant, OmpA, is thought to adhere to the cell-wall in the absence of Braun's Lipoprotein (lpp).

E. coli to osmotic stress [156] and observing how the cell envelope responded to the difference in turgor pressure across the envelope. Furthermore, they argue that across variations in osmalarity, the periplasm acts similar to the cytoplasm such that there is osmotic regulation occuring within the periplasmic space. This further suggests that the periplasmic space should be $\sim 70\%$ water by mass. Under this assumption, then an accurate model of the periplasm can be constructed from the relative abundances of periplasm proteins found in Table 4.4. The most abundant periplasm proteins tend to be regulatory proteins, or responsible for the capture of resources within the periplasm. For example, SodA, LivJ, and MglB, are all responsible for capturing superoxide dismutase, leucine/isoleucine/valine, and glucose/galactose respectively. Furthermore, a few other proteins such as HdeA/B respond to heat shock and Skp is a polypeptide chaperone. There is some liberty in choice of periplasm proteins depending upon the system of interest.

Lastly are the integral, membrane-associated, or lipo-proteins of the IM. First, it should be addressed that there are almost 1000 IM associated proteins, which does not include the cytoplasm facing IM associated proteins (> 500 more). However, constructing the IM is where flexibility comes in when constructing an envelope model. If the burden would be to construct an entire cell, then the information contained would need to be expanded greatly to accomodate. This is due to the fact that many of the proteins (or in the case of abundances, genes) localized to the IM are uncharacterized. Other proteins, such as the small mechanosenstive channel protien (MscS) has been previously characterized [157]. While this is simultaneously disheartening, it provides for great opportunity to generate all-atom systems with interesting protein-complexes such as the AcrA/B-TolC multidrug efflux complex, the TonB-ExbB/D system, the Lipopolysaccharide translational machinery, ATP synthase, etc.

4.4 Future Directions

The entire cell envelope of Gram-negative bacteria provides ample space for exploring the associated dynamics of cellular uptake, regulation, and excretion. The culmination of whole cell proteomics data as well as ongoing experiments to determine protein localizations, crystal structures, as well as protein-protein, protein-membrane, and protein-cell-wall interactions provide opportunity for computational simulation. More specifically, recent experiments with LC-MS/MS have generated high confidence protein-protein interactions within sub-cellular compartments. These protein-protein interaction maps, along with abundance data, can be used to create high resolution, all-atom, microcosm environments where all the proteins contained within are interacting with each other. This opens up the possibility of gaining deeper insight into envelope-spanning protein complex function. Further directions for the cell-envelope project are towards constructing these all-atom models to better understand how these spanning complexes handle stress response, how insertion of lipids or LPS into the OM occurs, and begin to understand how the proteins in the periplasm and at the OM operate in the absence of cellular energy.

Locus	Gene	WholeCell	Apex12	emPAI12	iBAQ12	GPM14	Krug13	Krug-emPAI	Avg (ppm)
					OM Lipoproteins (20 of 98)				
b1677	lpp	6110	2389	1147	1130	5252	13197	8190	5345
b4411	ecnB*	634	1041	278	0	1432	6068	1	1351
b1739	osmE	723	630	516	1389	2441	2369	208	1182
b1641	slyB*	887	2183	774	340	2024	1624	73	1129
b0453	ybaY*	399	1542	453	888	1270	1708	16	897
b1431	ydcL*	570	1834	1244	928	377	1124	16	870
b2809	ygdI*	664	581	378	604	608	2575	3	773
b0741	pal	376	449	66	31	962	2385	93	623
b1283	osmB	98	0	0	0	2817	379	1	471
b3506	slp*	302	509	168	193	957	527	5	380
b0434	yajG*	277	746	535	377	512	154	1	372
b2477	bamC	269	238	342	264	536	579	28	322
b4144	yjeI*	352	615	104	125	516	479	19	316
b2617	bamE	326	1003	104	0	212	514	2	309
b0865	ybjP*	177	127	105	155	411	1076	0	293
b2512	bamB	199	455	284	231	396	363	3	276
b1105	lpoB	128	613	119	130	276	138	5	201
b3150	yraP*	139	232	134	108	309	354	10	184
b2595	bamD	151	97	214	179	335	239	7	175
b1928	yedD*	149	125	135	65	359	178	22	148

Table 4.2: Twenty most abundant OM associated lipoproteins. Lpp, also known as Braun's lipoprotein, is primarily responsible for anchoring the cell wall to the OM. In the absence of Lpp, it has been shown [153] that OmpA can also bind and attach to the cell wall. Other lipoproteins such as ecnB and osmE are associated with bacterial funciton under osmotic shock. SlyB is associated with bacterial response to fluctutaions in $[Mg^{2+}]$. (*Crystal structures do not currently exist for these proteins.)

Locus	Gene	WholeCell	Apex12	emPAI12	iBAQ12	GPM14	Krug13	Krug-emPAI	Avg (ppm)
Periplasm facing OM associated (3 of 8)									
b0819	ybiS	348	958	545	561	506	644	11	510
b1583	ynfB	159	484	470	182	73	79	5	207
b1678	ynhG	104	271	70	90	151	270	6	137
Extra-cellular facing OM associated (5 of 19)									
b1677	lpp	6110	2389	1147	1130	5252	13197	8190	5345
b2779	eno	3361	5083	5060	6352	5709	4228	131	4275
b1923	fliC	1212	1949	2813	3106	705	1	-	1631
b1042	csgA	109	807	85	818	6	-	-	365
b4314	fimA	21	0	0	0	208	494	0	103
Excreted (3 of 6)									
b4376	osmY	463	853	672	1430	1813	93	1	761
b1071	flgM	251	994	512	184	43	-	-	397
b1847	yebF	277	708	813	805	164	1	0	395

Table 4.3: Top: Three of eight most abundant periplasm facing OM-associated proteins. These proteins differ from OM-lipoproteins in that they are not lipidated, and are either not considered $\beta-$barrel proteins, or do not span the membrane. Center: EC-facing OM associated proteins. Bottom: Proteins found to be excreted by the bacterium.

Locus	Gene	WholeCell	Apex12	emPAI12	iBAQ12	GPM14	Krug13	Krug-emPAI	Avg (ppm)
Periplasmic proteins (20 of 320)									
b3908	sodA†	4185	8335	12186	12390	4684	3576	407	6538
b3460	livJ†	2975	4888	9135	8250	3834	272	6	4194
b2150	mglB†	2392	3229	8520	5631	1665	91	1	3076
b1920	fliY	2574	2185	9879	2897	922	792	22	2753
b2310	argT†	2133	3042	7734	5214	642	70	2	2691
b0811	glnH†	2341	2729	8097	4165	1250	65	2	2664
b1857	znuA†	2600	0	0	0	1779	6694	6003	2439
b3544	dppA†	1455	2621	3760	5359	1474	1064	31	2252
b3510	hdeA	2045	598	935	7560	3948	6	0	2156
b3751	rbsB†	1392	2086	4410	3711	3169	193	5	2138
b0329	yahO†	1111	3154	2840	7005	654	63	0	2118
b1243	oppA†	1187	1942	1737	2930	1753	4591	68	2030
b4034	malE†	1392	3314	4124	3935	1009	6	0	1969
b0178	skp	796	1408	1283	2738	1913	1518	11	1381
b3509	hdeB	820	2634	1111	3196	1054	10	0	1261
b0655	gltI†	885	1775	2380	1470	1872	99	9	1213
b1973	zinT†	237	0	0	0	3992	4151	4	1198
b1440	ydcS	854	2292	2764	2133	186	26	3	1180
b0753	ybgS	606	1813	1139	1057	907	-	-	1104
b1123	potD†	561	1862	1059	1601	569	169	10	833

Table 4.4: Twenty of the most abundant periplasm proteins (not membrane associated). Majority of proteins shown here are capture proteins†; responsible for capturing free diffusing amino acids, oligopeptides, ions, or carbohydrates in the periplasm. fliY is a Flagellar motor assocaite protein, hdeA/B are heat-shock proteins and skp is a polypeptide chaperone for OM protein insertion.

Locus	Gene	WholeCell	Apex12	emPAI12	iBAQ12	GPM14	Krug13	Krug-emPAI	Avg (ppm)
					Integral IM (20 of 969)				
b3098	yqjD	1381	2150	2486	3267	1537	2820	4	1949
b2266	elaB	1009	1699	1729	1151	972	2276	24	1266
b0407	yajC	1595	780	423	1009	2180	1814	56	1122
b3736	atpF	701	963	946	751	1292	1365	161	883
b3233	yhcB	475	1137	565	399	593	869	27	581
b0570	cusS	72	0	0	0	3492	1	0	509
b2924	mscS	233	442	218	214	191	1845	31	453
b2092	gatC	269	1267	238	998	8	75	0	408
b3178	ftsH	326	346	274	349	496	934	92	402
b1101	ptsG	313	546	204	542	506	519	27	380
b1819	manZ	316	889	419	506	71	413	11	375
b0598	cstA	230	1511	152	488	194	14	0	370
b0733	cydA	614	736	160	418	360	285	3	368
b1602	pntB	288	340	163	425	496	721	19	350
b1737	chbC	33	0	0	0	1715	-	-	350
b4174	hflK	277	366	256	267	475	784	5	347
b0441	ppiD	249	415	400	287	406	606	8	339
b2027	wzzB	218	971	156	143	315	527	13	335
b1279	LapA	544	0	0	178	145	172	1239	325
b2513	yfgM	275	421	494	263	283	429	21	312

Table 4.5: Twenty of the most abundant integral IM proteins. Many of the most abundant IM proteins shown within this table vary in function. Some perform regulatory functions such as MscS, others such as Wzz are responsible for regulating the length of O-antigens expressed on LPS. Although there are over 900 localized to the IM, only ~ 100 have an abundance over 50 ppm across the seven abundance measurements shown, suggesting that either those proteins are expressed in low quantities or have not previously shown up in abundance measurements.

Locus	Gene	WholeCell	Apex12	emPAI12	iBAQ12	GPM14	Krug13	Krug-emPAI	Avg (ppm)
IM Lipoprotein (5 of 24)									
b0197	metQ	1314	2467	2185	3156	1694	4976	47	2263
b0463	acrA	256	391	236	435	437	887	8	379
b3554	yiaF	315	978	356	220	294	462	21	378
b3472	dcrB	262	310	355	324	359	774	88	353
b0432	cyoA	166	337	164	295	442	527	24	279
Periplasm facing, IM associated (5 of 22)									
b1614	ydgA	384	1120	731	555	366	777	45	568
b0839	dacC	249	409	557	604	456	606	11	413
b0632	dacA	195	309	287	207	698	585	24	329
b1830	prc	153	605	230	201	253	204	9	236
b0574	cusB	60	168	160	120	22	-	-	106

Table 4.6: Combination of IM lipoproteins and periplasm facing IM associated proteins. with the exception of the methionine binding protein MetQ, many of the proteins shown have relatively low abundance when compared to other groups. Of note is AcrA, one of the lipoproteins associated with the multidrug efflux pump.

CHAPTER 5
CONCLUSION

Over the past three decades, great strides have been made in developing encompassing bio-molecular force fields to describe atomistic behavior, compuational power has grown exponentially, and the amount of subcellular information has exploded. The culmination of this information has permitted computational dynamicists to use their computers as a molecular microscope to study the phenomena and behavior of proteins, membranes, and other biological molecules. Naturally, an area of study that has grown with the field of biophysics is at the cellular interface, where regulation and nutrient uptake are critical for cell vitality.

The first study, presented in chapter 2, simulated the outer-membrane transporter BtuB in a symmetric phospholipid membrane as well as an asymmetric LPS-containing membrane. Results from the simulations, published in Biophysical Journal, suggest that the OM environment plays a role on the function of OM associated proteins, which was also observed in another outer membrane porin [53]. Furthermore, the allosteric effect that the membrane composition has within the $\beta-$barrel effects the force required to extract the luminal domain from the barrel. In comparison to the Ca^{2+}-bound state, while the LPS oligosaccharides stabilize the extracellular loops of apo-state BtuB, the stabilization does not create an adequate binding site for B12, the substate that binds BtuB. Therefore, it is suggested from observed simulations that future simulations of OM proteins be performed in native lipid membranes to capture membrane-protein interactions. In order to further the study of TonB-dependent transporters (TBDTs) in their native environment, it is suggested to look at two different areas, the first being substrate binding at the extracellular interface, as well as luminal domain extraction. The first method could be directed such that the signaling mechanism for dislocating the Ton-box can be better understood and therefore

targeted for antibiotic uptake, which ties directly to the second method of luminal domain extraction. Previous studies have suggested mechanisms for partial or full luminal-domain extraction [18]. Since the luminal-domain of TBDTs is highly conserved, it would be possible to see if and how the signal transduction is maintained across homologous proteins.

Next, as the move to larger system sizes becomes more prevalent, methods to enhance the computational output while reducing computational costs are necessary. As outlined for proteins, by Hopkins *et al.*, hydrogen mass repartitioning (HMR) provides a means to speed up computational output while utilizing the same amount of resources. This is done by reweighting the mass of the hydrogens attached to heavy-mass parent atoms of proteins, while maintaining the mass of the molecule [59]. This method has been applied to multiple systems through an accelerated MD software with a 4 fs timestep and a short, 9 Å cutoff for long-range interactions. Furthermore, this method was being applied to membrane-protein systems without validation on various lipid types. Here, we present a validation of utilizing HMR with various time steps and cutoffs to determine the effect of HMR, timestep, and cutoff on the production of membrane systems. The results from pure membrane suggest that there is a negligible effect of HMR when comparing conventional MD systems. Second, the change in time-step incurs a small penalty when comparing lipid order parameters, and only a small change in the diffusion observed in simulation. It is the cutoff of the simulations that create the greatest change across all simulations performed. This is unsurprising as lipids primarily interact by London-dispersion and hard-sphere repulsion, governed by the Lennard-Jones potential, such that changes to the cutoff of the non-bonded LJ interaction will have an adverse effect on production. While a difference is expected, the results from HMR simulations suggest that using a 9 Å cutoff drastically effects structural and kinetic properties of lipids and should therefore be used cautiously, as the effects observed in membrane-protein simulations are not as egregious.

Lastly, building upon simulations in the OM and accelerations of simulation, we turn to the components necessary for an all-atom cell-envelope model for Gram-negative bac-

teria. As outlined, there are multiple databases of whole cell proteomics information with corresponding localizations and abundances. The gretaest limitation in constructing a cell-envelope model is in determining the pieces and if those exist. From the tables shown in Chapter 4, many of the abundant proteins at various cellular localizations have no structural information currently and are unable to be included in an envelope model. However, the abundance data as provided sheds light on relative measures and amounts of proteins to be included in smaller, simulation systems. In order to construct theses microsystems, more information about protein-protein interactions must be elucidated experimentally. Work is being done currently to determine these interactions with high confidence, such that cell-envelope interaction maps can be created and influence system construction. Therefore, it is suggested that taking in protein interaction data and supplement with protein abundances to generate envelope models to study envelope spanning complexes will be the next great leap in computational simulation.

CHAPTER 6

SUPPLEMENTAL INFORMATION: ROLE OF THE NATIVE OUTER-MEMBRANE ENVIRONMENT ON THE TRANSPORTER BTUB

The original supplemental information can be found in **Biophysical Journal**

Volume 111. Issue 7. p1409–1417. October 2016.

doi:https://doi.org/10.1016/j.bpj.2016.08.033

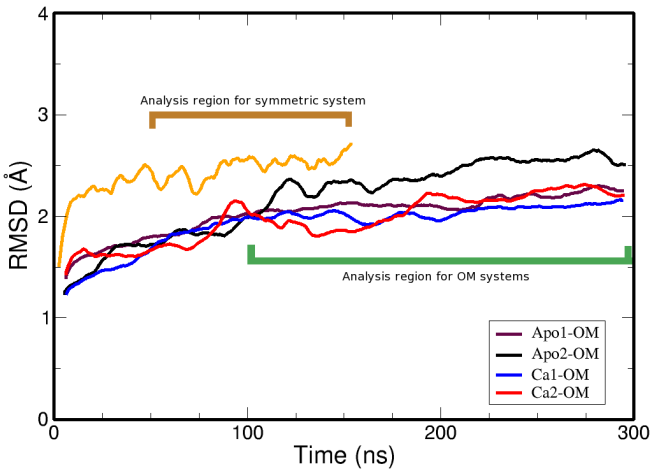
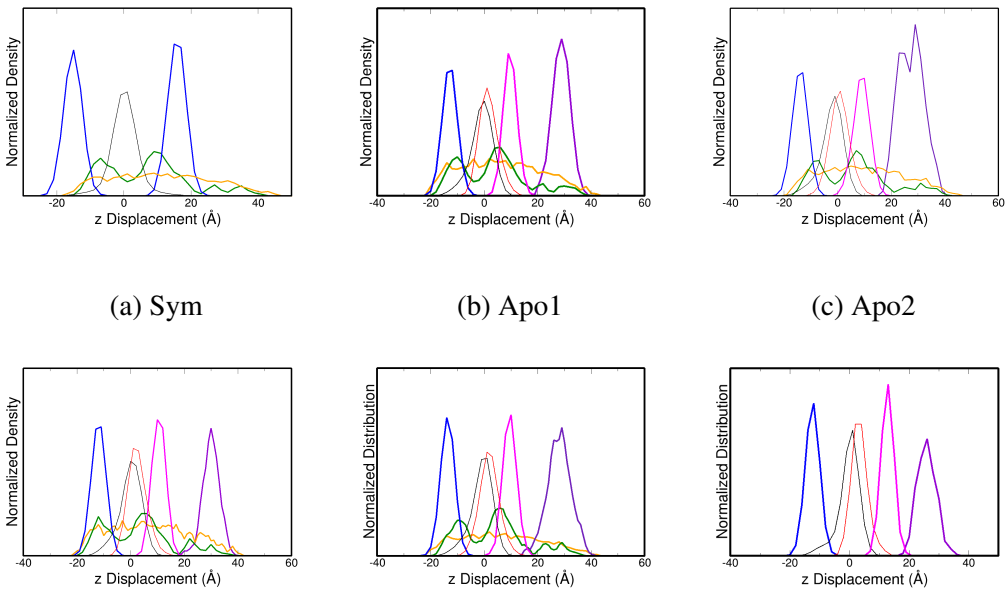


Figure 6.1: RMSD of the protein backbone for each simulated system compared to the CBL-bound crystal structure. Data regions used for analysis in each system has also been highlighted.



(a) Sym (b) Apo1 (c) Apo2

(d) Ca1 (e) Ca2 (f) OM

Figure 6.2: Plot of the hydrophobic thickness profiles for the simulated systems. In each system, the protein density from carbon$_\alpha$ is shown in orange, lipid A acyl C2 and C4 atoms in magenta, the core oligosaccharides (LPS headgroup) in purple, POPE acyl C2 carbons in blue and the aromatic protein sidechains are in green. The number density of the last carbon atoms of the POPE acyl tails (black) and the last carbons on each of the lipid A tails (red) represent aliphatic tail interdigition at 0 z-displacement. The hydrophobic thickness is determined between the blue and magenta curves.

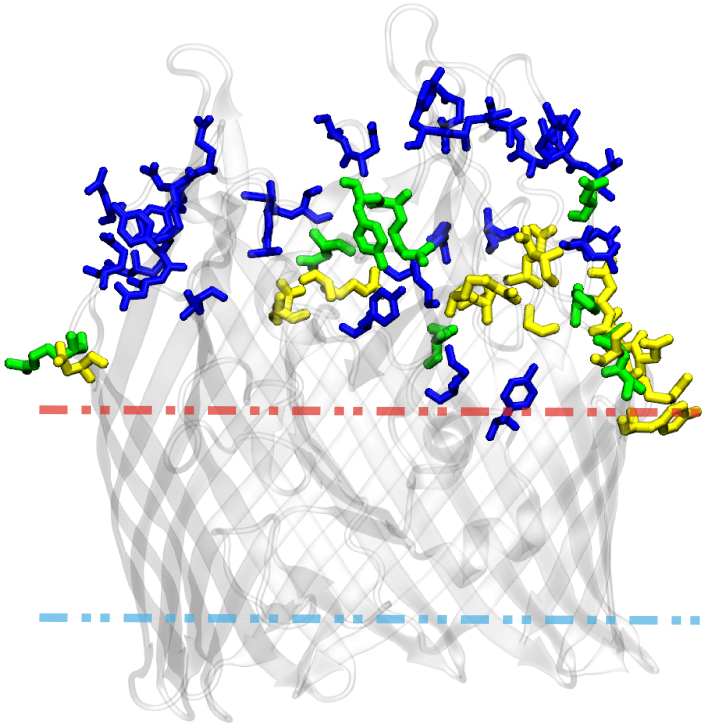
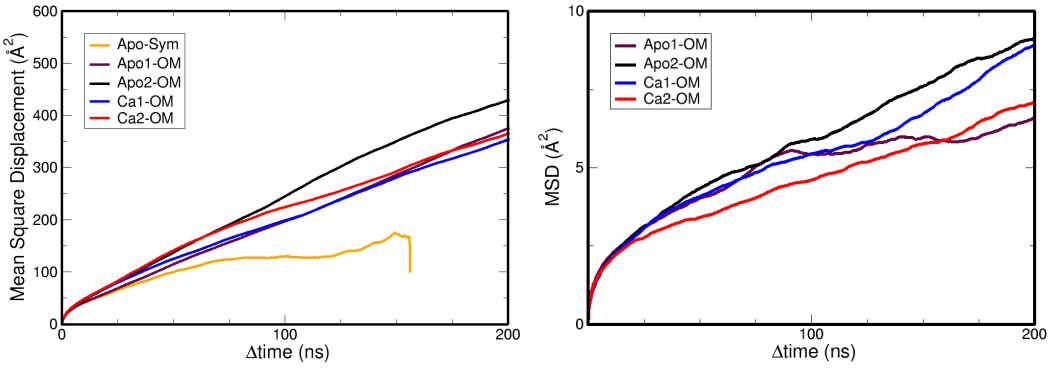


Figure 6.3: Representation of the amino acids that hydrogen bond with LPS (blue), POPE (yellow) or both (green). The hydrophobic matching region as determined from aromatic side chains is represented as red (extracellular) and blue (periplasmic) dashed lines. Side chains that interact with LPS tend to be much farther up the EC loops than those interacting with POPE.



(a) Phospholipid (b) LPS

Figure 6.4: Plots of Mean Square Displacement of (A) lipids in the inner phospholipid-containing leaflet and (B) outer LPS-containing leaflet. The diffusion coefficients for each leaflet are provided in Table S2.

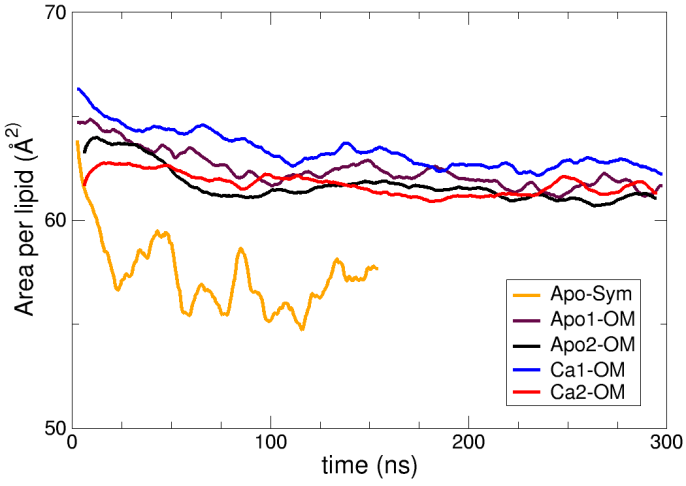


Figure 6.5: Plot of the area per lipid versus time for the inner phospholipid containing leaflet of each system. Each of the LPS-containing model membranes have an APL convergence around 63 Å^2.

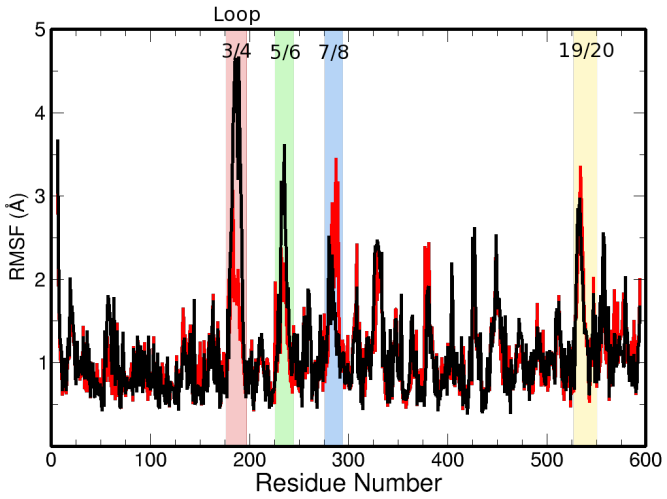


Figure 6.6: RMSF plots of Apo2-OM and Ca2-OM. Apo2-OM is shown in black and Ca2-OM is shown in red. Similar to the observations in Apo1- and Ca1-OM, there is a significant reduction in the fluctuation of loop 3/4 as well as a slight reduction in loop 5/6.

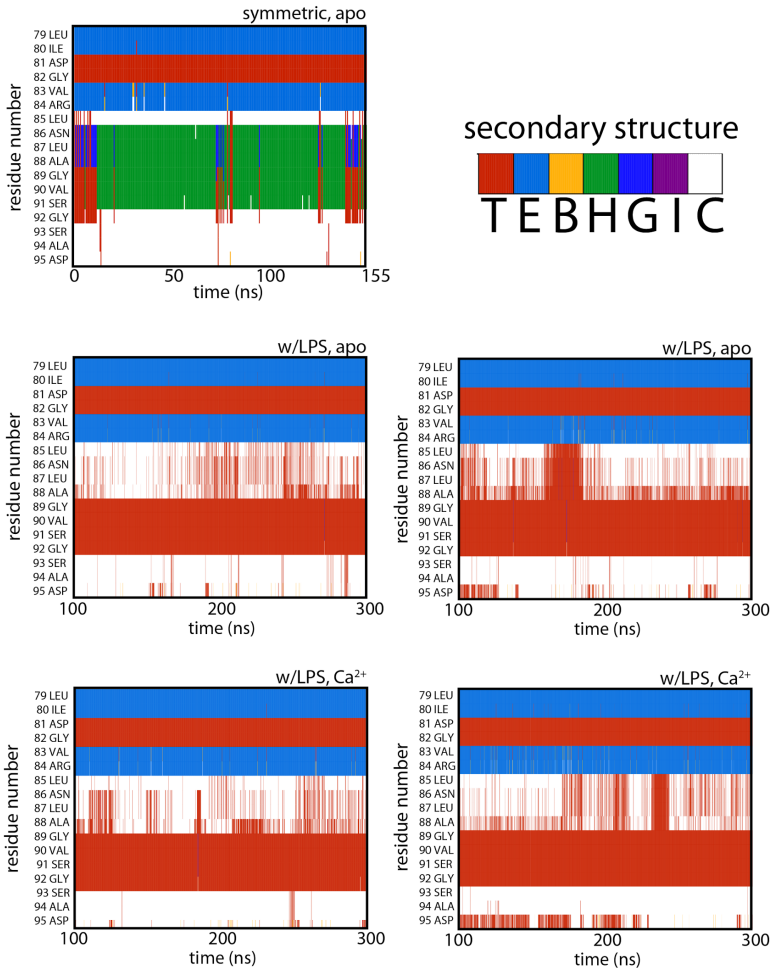
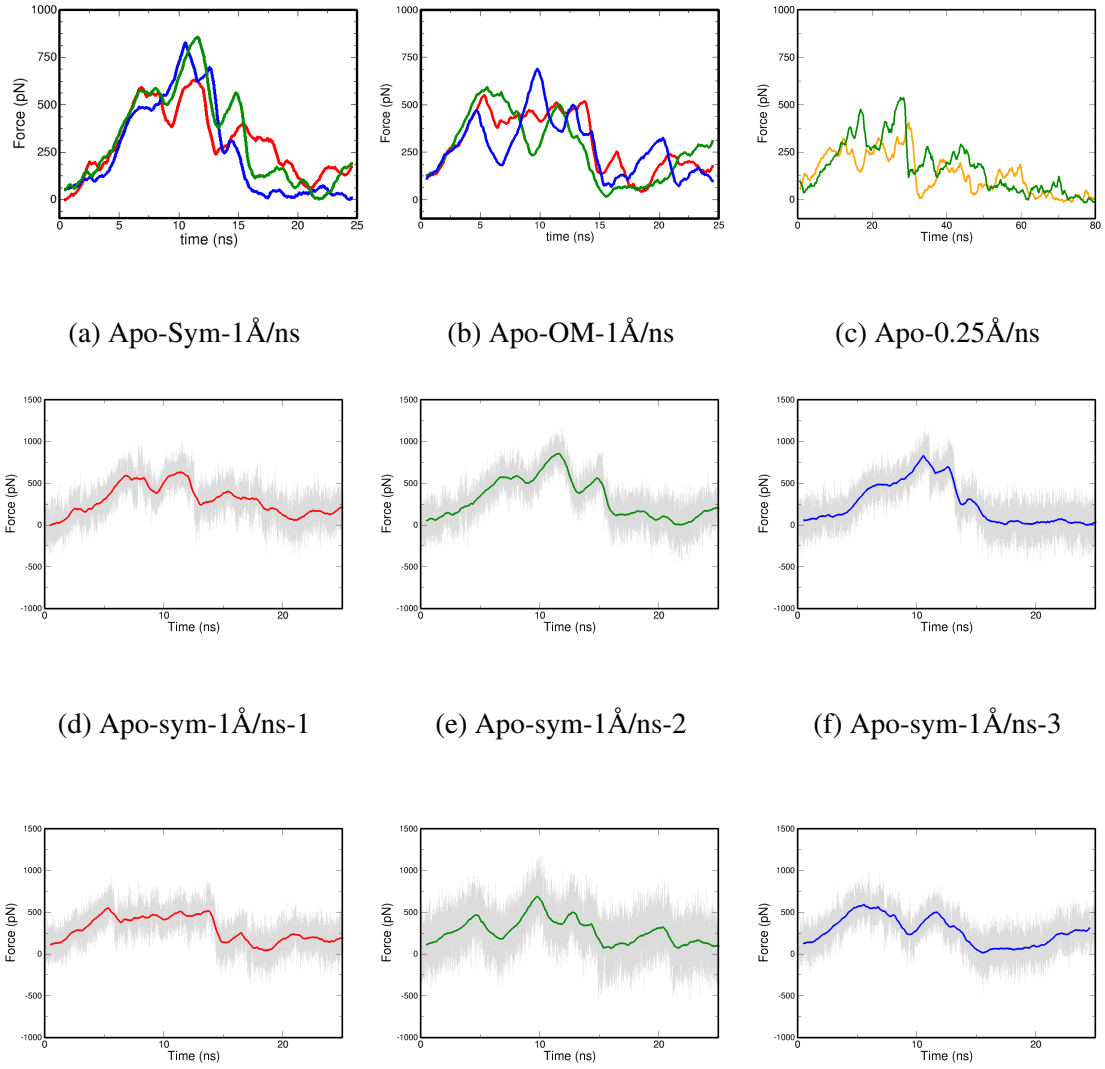


Figure 6.7: Timeline plots for the luminal domain region of Leu79 to Asp95. The $\alpha-$helix forms quickly in the Apo-sym simulation whereas the random coil is more prominent in the OM systems between residues Asn86 and Ser91. This random coil configuration permits Ser91 to be closer to the organization observed in the CBL-bound crystal structure.

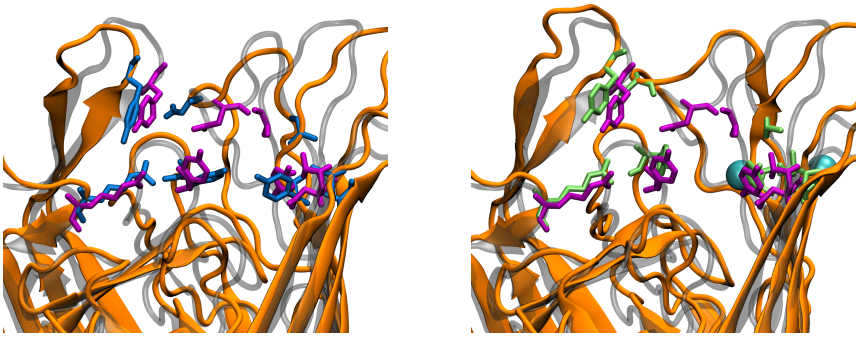


(a) Apo-Sym-1Å/ns (b) Apo-OM-1Å/ns (c) Apo-0.25Å/ns

(d) Apo-sym-1Å/ns-1 (e) Apo-sym-1Å/ns-2 (f) Apo-sym-1Å/ns-3

(g) Apo-OM-1Å/ns-1 (h) Apo-OM-1Å/ns-2 (i) Apo-OM-1Å/ns-3

Figure 6.8: Time averaged plots from the SMD extraction of the Ton box. (A) Three separate SMD simulations at 1 Å/ns of luminal domain extraction in the Apo-sym simulation. Start frames were seeded from 100 ns (red), 100 ns (blue), and 150 ns (green). (B) Three separate SMD simulations at 1 Å/ns of luminal extraction in the Apo1-OM (red, blue) and Apo2-OM (green) simulation. Start frames were seeded from Apo1-OM at 100 ns (red), 150 ns (blue), and Apo2-OM at 100 ns (green). (C) Force curve from constant SMD lumen extraction performed at 0.25 Å/ns. The Apo-OM system (orange) produces a smaller extraction force compared to the Apo-sym system (green). (D-I) Plots of time averaged curves shown in (A) and (B) with accompanying raw data.



(a) Apo-OM2 (b) Ca-OM2

Figure 6.9: Comparison of the (A) Apo- and (B) Ca^{2+}-bound states of BtuB in the second simulated OM systems. (A) Apo-state BtuB (orange ribbon) aligned to the CBL-bound crystal structure (transparent grey ribbon) from Chimento et al. [83]. Residues within 4 Å of CBL in the CBL-bound crystal structure are highlighted as blue (Apo2-OM) and purple (crystal structure) sticks. (B) Ca^{2+}-bound BtuB (orange ribbon) aligned to the CBL-bound crystal structure (transparent grey ribbon). Calcium ions are shown as cyan spheres and the green sticks represent the CBL binding residues in the Ca2-OM system.

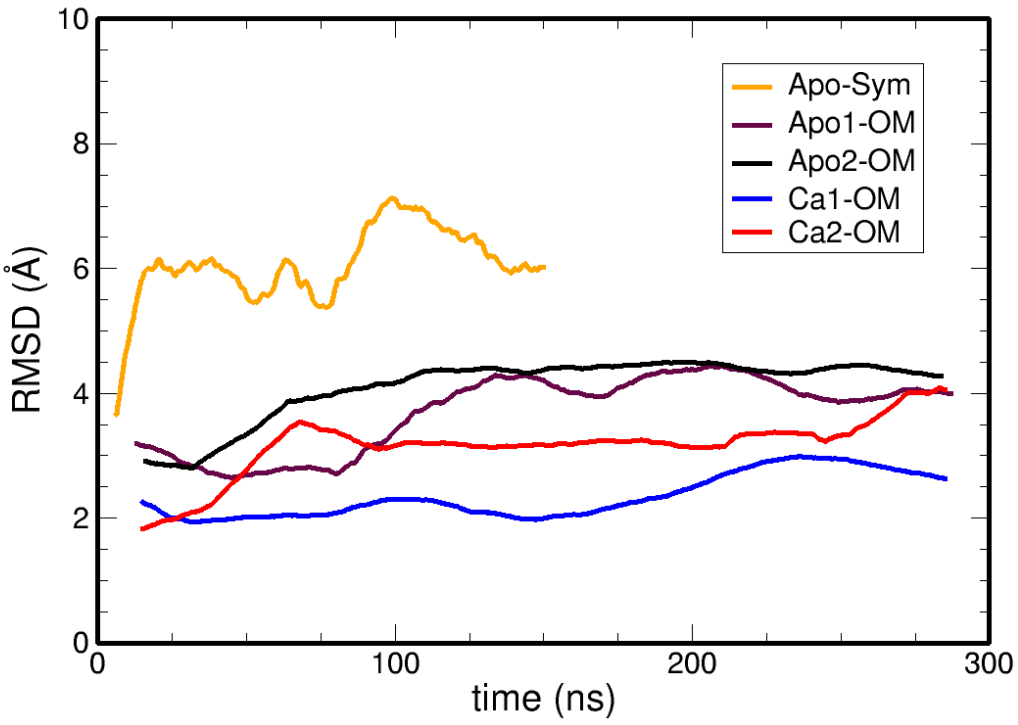


Figure 6.10: RMSD of Loop 3/4 in each of the simulated systems. RMSD was calculated for the sidechains and backbone of His176 to Lys200 in comparison to the CBL-bound crystal structure. Greatest amount of deviation (> 8Å) is observed in the Apo-sym system with decreased values for the Apo1- and Apo2- as well as Ca1- and Ca2-OM.

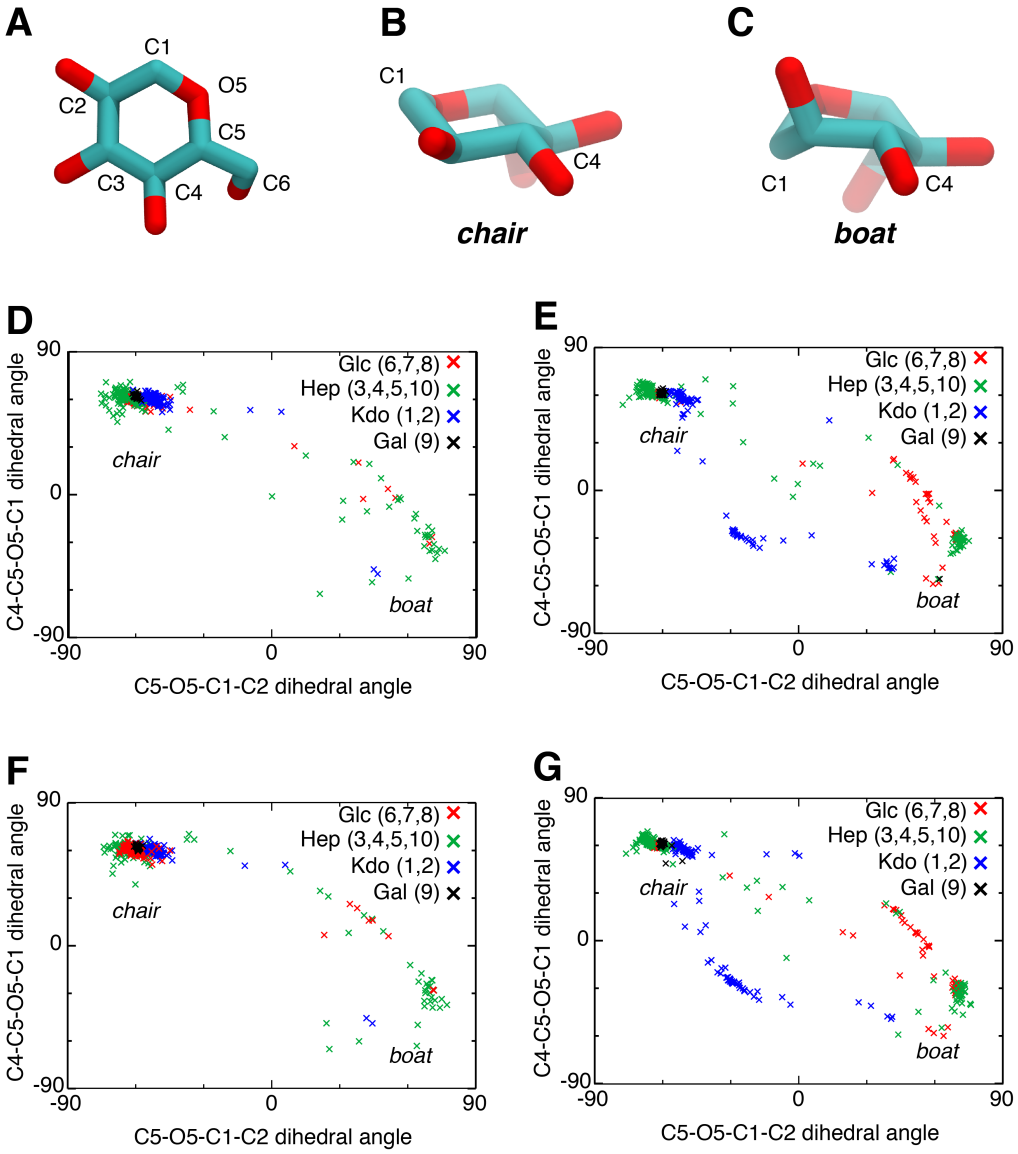


Figure 6.11: (A) Top down view of a pyranose ring, found in each of the sugars in the core oligosaccharide. The (B) chair conformation of a pyranose ring is the energetically favored conformation as compared to the (C) boat conformation. Determination of sugar conformation was done by measuring the dihedral angels on either side of the O5 oxygen. (D-G) include results from each of the OM simulations as an average dihedral conformation throughout the trajectory.

Table 6.1: Summary of membrane results for each BtuB system.

System (time ns)	Bilayer composition (Top:Bottom)	Diffusion Coefficient (cm^2/s)	Area Per Lipid ($Å^2$/ Molecule) $\pm$ Std. Dev.	Hydrophobic Thickness (Å)
Apo-sym (150 ns)	66-POPE 61-POPE	2.3×10^{-8} 2.0×10^{-8}	58.8 ± 0.4 63.7 ± 1.3	34
Apo1-OM (300 ns)	50-LPS 157-POPE	7.2×10^{-10} 4.4×10^{-8}	194.6 ± 1.6 61.9 ± 0.5	24
Apo2-OM (300 ns)	50-LPS 157-POPE	6.2×10^{-10} 4.3×10^{-8}	192.6 ± 1.3 61.4 ± 0.4	24
Ca1-OM (300 ns)	50-LPS 157-POPE	9.8×10^{-10} 4.5×10^{-8}	197.3 ± 1.3 62.8 ± 0.4	26
Ca2-OM (300 ns)	50-LPS 157-POPE	6.0×10^{-10} 4.6×10^{-8}	193.1 ± 1.5 61.5 ± 0.5	25
OM only (172 ns)	36-LPS 102-POPE	8.1×10^{-10} 6.4×10^{-8}	178.9 ± 0.6 63.1 ± 1.4	25

Table 6.2: Summary of hydrogen bonds from the Apo1-OM and Apo2-OM systems. Occupancies over 100% are due to multiple hydrogen bonds forming simultaneously.

BtuB Residue	LPS component	Apo1-OM occ.	Apo2-OM occ.
Loop 1/2			
SER146	LipidA	91.75	-
ASN147	LipidA	96.30	-
Loop 3/4			
TYR172	LipidA	-	11.10
ASP179	Gal	11.10	-
ASP179	Gal	21.80	-
ASP179	Gal	38.55	-
ASP179	Glc	124.85	-
ASP179	KDO	-	13.95
TYR183	KDO	-	25.15
TYR183	KDO	-	10.50
GLY184	Hep	-	21.10
THR186	Hep	-	22.05
THR188	Hep	-	11.45
THR192	Hep	17.00	-
ASP193	Hep	100.30	-
ASP193	Gal	20.80	15.15
ASP193	Glc	18.85	-
ASP193	Glc	15.30	-
ASN194	Hep	17.35	-

Continued on next page

Table S6.2 – continued from previous page

BtuB Residue	LPS component	Apo1-OM occ.	Apo2-OM occ.
ASN194	Hep	24.10	-
ASN194	Gal	15.00	-
ASN194	Gal	19.20	-
LYS200	LipidA	-	73.05
Loop 5/6			
ASP241	Glc	23.70	-
THR242	Glc	10.70	-
LYS244	LipidA	-	33.30
TYR246	LipidA	-	53.35
GLN248	LipidA	29.40	-
TRP250	LipidA	-	10.70
Loop 7/8			
HSE280	Hep	11.00	-
TYR281	Hep	11.25	-
ARG283	KDO	-	11.15
ARG283	LipidA	-	14.95
TYR284	Hep	-	28.05
ASP285	Hep	218.50	35.25
ASP285	Hep	61.25	
SER286	Hep	19.65	-
SER286	Hep	44.65	-
SER287	Glc	-	11.00
SER287	Glc	-	12.80
Loop 9/10			
THR321	KDO	-	26.70
THR321	KDO	-	61.00
GLU330	Glc	16.35	-
GLU330	Gal	54.15	-
ASP331	Glc	35.15	85.40
ASP331	Glc	36.00	
ASP331	Hep	27.60	-
GLY332	Hep	66.10	-
GLY332	Gal	-	16.30
GLY332	Glc	-	38.20
TYR333	Glc	19.40	10.45
TYR333	Glc	24.30	
TYR333	Hep	-	23.95
TYR333	Hep	-	14.15
TYR333	Gal	-	11.20
ASP334	Gal	-	86.10
Loop 11/12			
GLN364	Gal	24.05	24.70

Continued on next page

Table S6.2 – continued from previous page

BtuB Residue	LPS component	Apo1-OM occ.	Apo2-OM occ.
ARG367	KDO	76.65	-
ARG367	LipidA	58.20	189.05
Loop 13/14			
TYR402	Gal	25.55	22.25
TYR402	Gal	11.55	
PRO408	Glc	-	13.40
ASN409	Gal	17.35	-
ASN409	Gal	10.75	-
ASN409	Glc	-	57.15
ASN409	Glc	-	50.15
ASP411	Hep	15.75	27.15
ASP411	Glc	65.75	-
Loop 15/16			
ASP442	Hep	117.45	-
GLU456	Gal	51.80	-
GLU456	Gal	22.75	-
ARG460	Hep	17.65	-
Loop 19/20			
THR525	LipidA	122.10	-
TYR527	KDO	51.70	-
TYR527	Hep	-	116.35
GLN537	Glc	10.70	13.20
GLN537	Glc	22.10	19.10
GLN537	Hep	29.15	-
THR538	Glc	37.45	-
LYS540	LipidA	105.15	-
LYS540	Glc	16.85	-
LYS540	KDO	36.85	-
LYS540	Hep	17.25	235.75
Loop 21/22			
LEU568	LipidA	-	19.20
ASP570	LipidA	-	146.15
LYS571	LipidA	34.50	191.10
TYR577	Hep	-	42.75
GLN580	Gal	10.65	-
GLN580	Gal	10.95	-
ARG584	LipidA	20.70	-
ARG584	KDO	-	327.05
TYR586	LipidA	28.10	-

Table 6.3: Summary of hydrogen bonds from the Ca1-OM and Ca2-OM systems. Occupancies over 100% are due to multiple hydrogen bonds forming simultaneously.

BtuB Residue	LPS component	Ca1-OM occ.	Ca2-OM occ.
Loop 1/2			
SER146	LipidA	-	11.33
Loop 3/4			
HSE174	LipidA	-	102.20
TYR183	Gal	17.65	-
ALA190	Gal	39.40	-
GLN191	Hep	-	15.34
GLN191	Gal	12.70	-
THR192	Hep	36.50	10.94
THR192	Hep	12.55	
THR192	Gal	10.60	-
ASP193	Hep	-	93.06
ASP193	Gal	12.55	-
ASN194	Hep	18.95	-
ASN194	Hep	14.75	-
ASN194	Glc	17.00	-
ASN194	Glc	46.75	-
Loop 5/6			
ARG226	LipidA	-	148.38
TYR246	LipidA	20.95	18.51
GLN248	LipidA	23.70	13.70
GLN248	LipidA	12.15	
Loop 7/8			
ARG283	LipidA	157.75	-
TYR284	KDO	-	50.19
TYR284	KDO	-	15.48
TYR284	LipidA	16.85	-
TYR284	Hep	-	19.97
ASP285	Hep	66.20	-
ASP285	Glc	27.45	-
SER286	Hep	48.30	-
SER286	Hep	34.85	-
SER286	Glc	22.10	-
SER286	Glc	14.05	-
Loop 9/10			
GLU330	Hep	11.35	-
GLU330	Hep	12.05	-
ASP331	Glc	69.40	-
ASP331	Hep	-	11.15
ASP331	Hep	-	54.62

Continued on next page

Table S6.3 – continued from previous page

BtuB Residue	LPS component	Ca1-OM occ.	Ca2-OM occ.
GLY332	Glc	19.65	-
TYR333	Glc	40.00	-
TYR333	Glc	29.55	-
TYR333	Gal	-	14.43
ASP334	KDO	88.15	-
ASP334	Hep	41.55	-
ASP334	Hep	19.55	-
Loop 11/12			
GLN364	Gal	-	13.07
GLN364	Glc	-	19.66
PHE365	Hep	-	14.88
ARG367	KDO	25.70	66.71
ARG367	LipidA	177.50	107.88
Loop 13/14			
TYR402	Glc	19.70	-
TYR402	Glc	26.80	-
ASN409	Glc	-	17.15
ASN409	Glc	-	23.35
Loop 15/16			
SER441	Hep	-	37.05
ASP442	Hep	-	50.26
LYS458	Hep	85.50	84.00
ARG460	Hep	-	197.56
Loop 17/18			
ARG487	KDO	55.30	-
ARG487	Hep	12.95	125.13
ASP492	Hep	14.50	-
LYS500	LipidA	19.25	-
Loop 19/20			
TYR527	Glc	15.45	-
GLN537	Hep	14.25	19.52
THR538	Hep	15.60	51.03
THR538	Glc	13.15	-
LYS540	Hep	16.55	139.94
Loop 21/22			
ASP570	Hep	-	67.41
LYS571	LipidA	10.40	207.39
LYS571	KDO	16.00	-
TYR577	Hep	-	54.41
GLN580	Hep	10.60	11.99
GLN580	Hep	16.90	
THR581	Hep	-	11.89

Continued on next page

Table S6.3 – continued from previous page

BtuB Residue	LPS component	Ca1-OM occ.	Ca2-OM occ.
ARG584	LipidA	167.90	84.07
TYR586	LipidA	30.90	56.54

CHAPTER 7

SUPPLEMENTAL INFORMATION FOR: HYDROGEN MASS REPARTITIONING FOR CHARMM36 MEMBRANE SYSTEMS IN NAMD

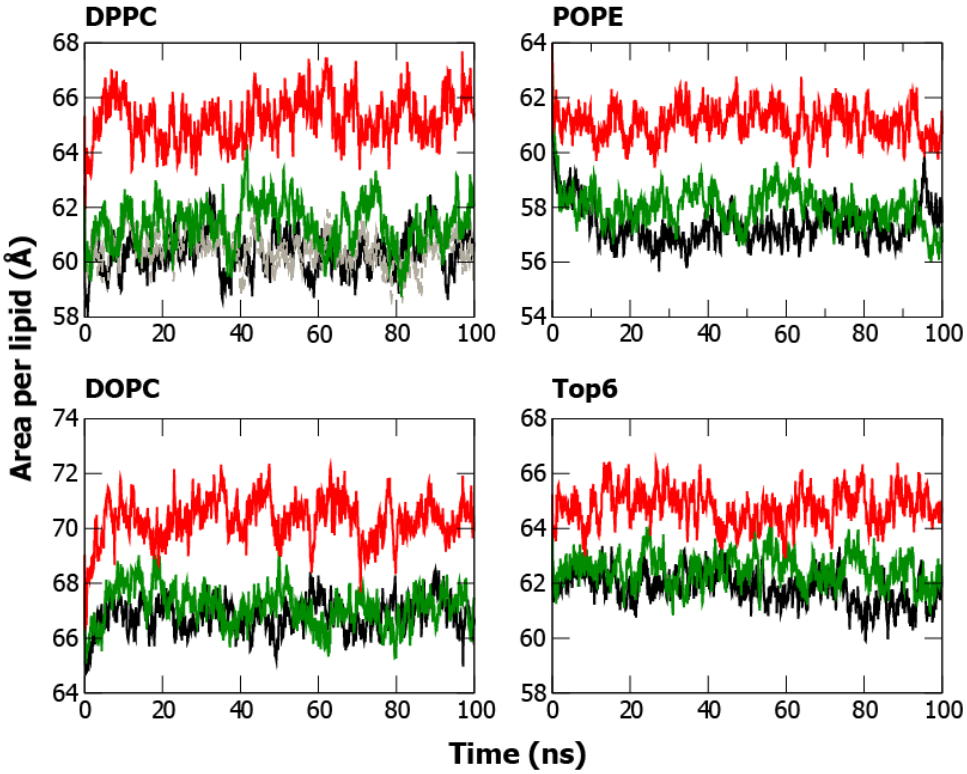


Figure 7.1: Plots of Area per lipid over 100 ns simulation period. Results presented for analysis were obtained from the last 50 ns of each simulation. In each panel, 2-12 in shown black, 4-12 in green, and 4-9 in red. A separate 2-12-HMR simulation performed in the DPPC (top left) bilayer is shown in grey.

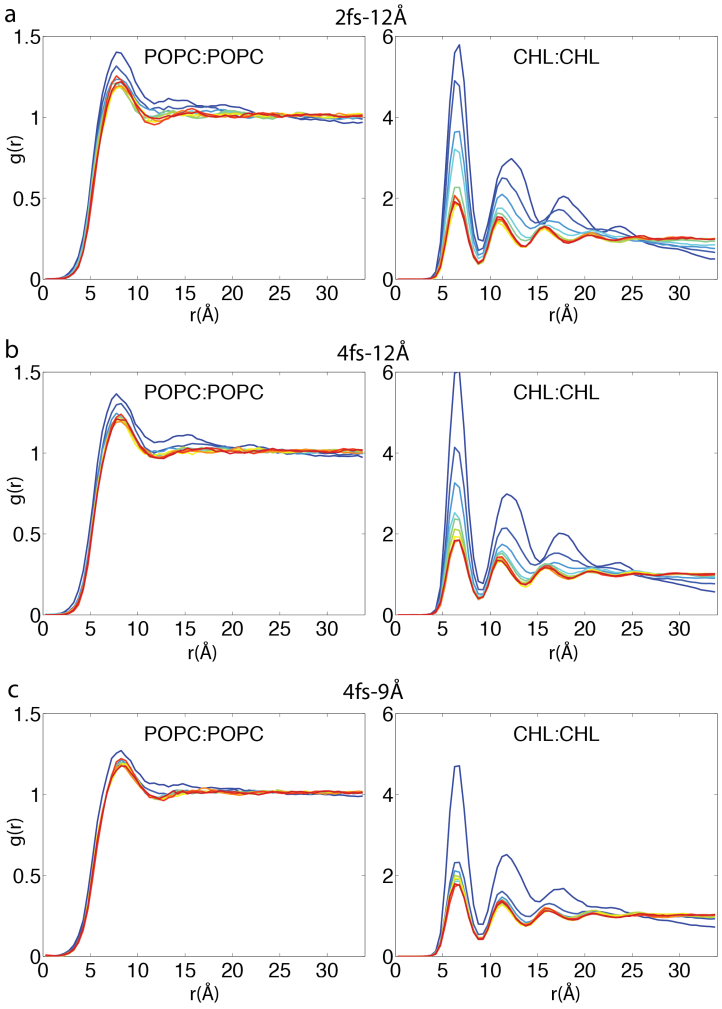


Figure 7.2: Time evolution of the radial pair distribution functions g(r) for POPC:POPC and CHL:CHL from the 2-12 simulation (a), the 4-12 simulation (b), and the 4-9 simulation (c). Calculated g(r) is averaged in 100-ns blocks and colored by simulation time, with blue, green and red indicating the beginning, the middle and the end of a simulation, respectively.

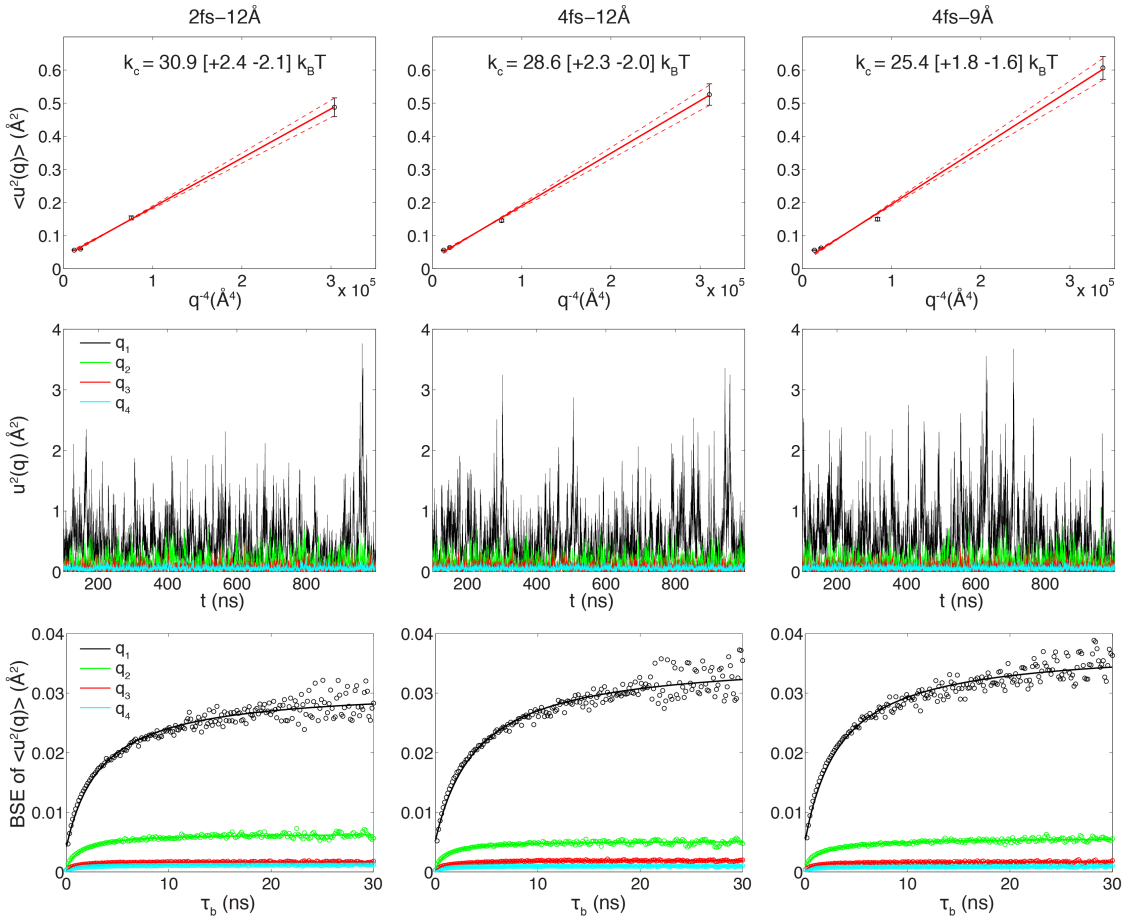


Figure 7.3: Undulation analysis of a POPC bilayer with 680 lipids. (a) The bending modulus k_c obtained from the slope of $\langle u^2(q)\rangle$ versus q^{-4}. The estimated uncertainty in k_c is given in square brackets. (b) $u^2(q)$ as a function of simulation time. (c) The blocked standard error (BSE) in $\langle u^2(q)\rangle$ as a function of the block size τ_b. Results of the lowest wavenumber (q_1 $\approx$0.04 Å^{-1}) are colored in black circles, with the black solid curve representing a least square fit using a rational polynomial function. Similarly, results for the next three wavenumbers are colored in green, red and blue, respectively.

System	Lipid Top:Bottom	Atom Count	H_2O/Lipid	Ions
DPPC	240:240	113064	35.125	84
POPE	240:240	107969	33.25	80
DOPC	240:240	122505	39.0	96
Top6	PMPE 116:116 POPE 32:32 QMPE 32:32 PMPG 24:24 PYPG 20:20 OYPE 16:16	109605	35.26	176
Membrane Mixing	POPC 70 top:bot CHL 35 top:bot	atoms	water:lip	ions
OmpF	POPE 88:88	99154	115.84	819
L8	POPE 53:52	26696	39.57	0

Table 7.1: Summary of systems performed for HMR. Pure membrane systems: DPPC, POPE, DOPC, Top6, and POPC:CHL membrane mixing. OmpF and L8 are membrane-protein systems.

Appendices

APPENDIX A

COMPUTATIONAL METHODS

A.1 Molecular Dynamics

The primary method utilized in performing computational simulations was through Molecular Dynamics (MD). Molecular dynamics simulation requires four primary components: molecular structure files, force-field parameters, classical mechanics energy calculation, and an iterative scheme to update atom positions. Each of these will be discussed in general detail below.

The first requisite component for MD simulation is an equation of motion, described by Newton's 2^{nd} law:

$$\vec{F} = m\frac{d^2\vec{x(t)}}{dt^2} = -\nabla U \tag{A.1}$$

Here we can approximate the potential energy of each atom within the system as a linear combination of harmonic oscillators:

$$\begin{aligned} U = & \sum_{bonds} k_b(b-b_o)^2 + \sum_{angles} k_\theta(\theta-\theta_0)^2 \\ & + \sum_{dihed} k_\phi[1+\cos(n\phi-\delta)] + \sum_{impr} k_\omega(\omega-\omega_0)^2 \\ & + \sum_{nonbond} \epsilon\left[\left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{12} - \left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{6}\right] + \sum_{elec}\frac{q_i q_j}{\epsilon r_{ij}} \end{aligned} \tag{A.2}$$

with bonded (e.g. bonds, angles, dihedrals, and impropers) and non-bonded (e.g. London dispersion, hard-sphere repulsion, and electrostatic) terms. The bonded terms are calculated at each time step for each atom within the simulation. Special care must be taken to appropriately calculate the non-bonded terms. First, the hard-sphere repulsion and London-dispersion interactions are captured by the Lennard-Jones potential (i.e. $(1/r)^{12} - (1/r)^6$

term) and truncated with a cutoff value. This cutoff is typically chosen such that the energy terms are approximately zero (around 12 Å). In order to smooth out the energy cutoff, a switching function is used to have the LJ energy go to zero at the cutoff. Similarly, electrostatic interactions are calculated using a Particle Mesh Ewald summation for calculating electrostatics across periodic boundaries.

Second, the potential energy calculation needs a defined parameter space in order to calculate the bonded and non-bonded energy terms, which are aptly-named force field parameters. These parameters are determined through quantum mechanics simulations and density functional theory approximations. The force fields utilized throughout this work were the CHARMM36 all-atom force fields for lipids, proteins, nucleic acids, carbohydrates, as well as the charmm general force field. Some parameterization was performed on the lipopolysaccharide model used in Chapter 2 and can be found in the supplement of the corresponding publication.

Third, it is necessary to have an iterative time-stepping scheme to update positions of particles within the simulation. This is done within NAMD [66] by using the verlet algorithm via the Störmer Method:

$$\vec{x}_{n+1} = 2\vec{x}_n - \vec{x}_{n-1} + \vec{a}_n \Delta t^2 \vec{a}_n = \vec{A}(\vec{x}_n) = \frac{\vec{F}(\vec{x}_n)}{m} \tag{A.3}$$

such that the positions of each atom can be updated at each time step without having to explicitly calculate the atom velocities. Other methods, such as Runga-Kutta are not utilized because they don't conserve the Hamiltonian, or like the Euler method, introduce greater error.

Lastly, but just as importantly, all-atom simulation requires structural information of small molecules. These are commonly found on various databases (e.g. RCSB.org) and provided from multiple different experimental datasets. Most crystal values are determined using x-ray crystallography in combination with cryo-electron microscopy (for multipro-

tein complex structures) or nuclear magnetic resonance (NMR) imaging. These structure files provide a list of atoms and their corresponding vector positions in three-dimensions. A structure is read in and visualized using VMD [109], and a structure file is compiled with associated CHARMM topology files. Topology files inform the adept computational biophysicist which atoms are bonded within the constructed system. This would include unresolved atoms from a crystal structure (such as hydrogens) due to poor resolution of the original structure.

A.2 Steered Molecular Dynamics

Steered Molecular Dynamics (SMD), is a method in which the constructed system is biased from an external force. This is done by attaching an imaginary particle to an atom or molecule of interest within the constructed and quilibrated system, by an imaginary spring. The imaginary particle is then biased by a constant force or velocity vector. Because this imaginary particle is attached to an atom or molecule by a spring, there is an associated response by the molecule to being biased. This response is then measured within the standard output an analyzed.

A.3 Hydrogen Mass Repartitioning

Hydrogen Mass Repartitioning (HMR) is utilized in order to increase the effective time step in the verlet integrator calculation. This means that the hydrogen masses associated with lipids or proteins are all increased by some fixed value by decreasing thir bonded parent atom by that same amount. For example, a methyl carbon (CH_3) has a mass of 12.010 amu and each hydrogen is 1.008 amu. When applying HMR, each of the hydrogens increase to 3.024 amu (increase mass by factor of 3) and the carbon would decrease to 5.962 amu $(12.010 - 3 \times 2.016)$. This results in a slower oscillation frequency for the angle and dihedral harmonics used in the previously described potential energy, such that a longer time step (maximum of 4 fs) can be used. The use of HMR has only been validated for

proteins and lipids. Variations of HMR have been performed on water previously with a resultant shift in various properties in the water model used.

REFERENCES

[1] J. Gumbart, M. Wiener, and E. Tajkhorshid, "Mechanics of force propagation in the TonB-dependent outer membrane transport," *Biophys. J.*, vol. 93, pp. 496–504, 2 2007.

[2] C. Bradbeer and A. Gudmundsdottir, "Interdependence of calcium and cobalamin binding by wild type and mutant BtuB protein in the outer membrane of *Escherichia coli*," *J. Bacteriol.*, vol. 172, pp. 4919–4926, 9 1990.

[3] Centers for Disease Control and Prevention, "Cephalosporin susceptibility among *Neisseria gonorrhoeae* isolates," *Morbidity and Mortality Weekly Report*, vol. 60, pp. 873–877, 2011.

[4] C. R. H. Raetz, "Biochemistry of endotoxins," *Annu. Rev. Biochem.*, vol. 59, pp. 129–170, 1990.

[5] J. F. Nagle and S. Tristram-Nagle, "Structure of lipid bilayers," *BBARM*, vol. 1469, no. 3, pp. 159–195, 2000.

[6] H. Nikaido, "Molecular basis of bacterial outer membrane permeability revisited," *Microbiol. Mol. Biol. Rev.*, vol. 67, no. 4, pp. 624–640, 2003.

[7] R. M. Shroll and T. P. Straatsma, "Molecular structure of the outer bacterial membrane of *Pseudomonas aeruginosa* via classical simulation," *Biopolymers*, vol. 65, pp. 395–407, 2002.

[8] T. Piggot, D. Holdbrook, and S. Khalid, "Electroporation of the *E. coli* and *S. aureus* membranes: Molecular dynamics simulations of complex bacterial membranes," *J. Phys. Chem. B*, vol. 115, pp. 13 381–13 388, 2011.

[9] E. L. Wu, O. Engstrm, S. Jo, D. Stuhlsatz, M. S. Yeom, J. B. Klauda, G. Widmalm, and W. Im, "Molecular dynamics and NMR spectroscopy studies of E. coli lipopolysaccharide structure and dynamics," *Biophys. J.*, vol. 105, pp. 1444–1455, 2013.

[10] C. R. H. Raetz and C. Whitfield, "Lipopolysaccharide endotoxins," *Annu. Rev. Biochem.*, vol. 71, pp. 635–700, 2002.

[11] P. F. Mühlradt, J. Menzel, J. R. Goldecki, and V. Speth, "Lateral mobility and surface density of lipopolysaccharide in the outer membrane of *Salmonella typhimurium*," *Eur. J. Biochem.*, vol. 43, pp. 533–539, 1974.

[12] R. Hiruma, A. Yamaguchi, and T. Sawai, "The effect of lipopolysaccharide on lipid bilayer permeability of $\beta-$lactam antibiotics," *FEBS Lett.*, vol. 170, no. 2, pp. 268–272, 1984.

[13] H. Nikaido and M. Vaara, "Molecular basis of bacterial outer membrane permeability," *Microbiol. Rev.*, vol. 49, pp. 1–32, 1985.

[14] W. Vollmer, "Bacterial outer membrane evolution via sporulation?" *Nat. Chem. Biol.*, vol. 8, pp. 14–18, 1 2012.

[15] T. J. Piggot, D. A. Holdbrook, and S. Khalid, "Conformational dynamics and membrane interactions of the *E. coli* outer membrane protein FecA: A molecular dynamics simulation study," *Biochim. Biophys. Acta Biomembr.*, vol. 1828, pp. 284–293, 2013.

[16] J. D. Faraldo-Gòmez, G. R. Smith, and M. S. Sansom, "Molecular dynamics simulations of the bacterial outer membrane protein FhuA: A comparative study of the ferrichrome-free and bound states," *Biophys. J.*, vol. 85, pp. 1406–1420, 2003.

[17] D. P. Chimento, R. J. Kadner, and M. C. Wiener, "Comparative structural analysis of TonB-dependent outer membrane transporters: Implications for the transport cycle," *Proteins: Struct., Func., Bioinf.*, vol. 59, pp. 240–251, 2005.

[18] J. C. Gumbart, M. C. Wiener, and E. Tajkhorshid, "Coupling of calcium and substrate binding through loop alignment in the outer membrane transporter BtuB," *J. Mol. Biol.*, vol. 393, pp. 1129–1142, 2009.

[19] E. Eren, J. Parkin, A. Adelanwa, B. Cheneke, L. Movileanu, S. Khalid, and B. van den Berg, "Toward understanding the outer membrane uptake of small molecules by *Pseudomonas aeruginosa*," *J. Biol. Chem.*, vol. 288, pp. 12 042–12 053, 17 2013.

[20] C. Wülfing and A. Plũckthun, "Protein folding in the periplasm of *Escherichia coli*," *Mol. Microbiol.*, vol. 12, no. 5, pp. 685–692, 1994.

[21] H. Nikaido, "Transport across the bacterial outer membrane," *J. Bioenerg. Biomembr.*, vol. 25, no. 6, pp. 581–589, 1993.

[22] V. Braun, C. Bös, M. Braun, and H. Killman, "Outer membrane channels and active transporters for the uptake of antibiotics," *J. Inf. Diseases*, vol. 183, S12–S16, 2001.

[23] N. Noinaj, M. Guillier, T. Barnard, and S. Buchanan, "TonB-dependent transporters: Regulation, structure and function," *Annu. Rev. Microbiol.*, vol. 64, pp. 43–60, 2010.

[24] N. Noinaj, N. C. Easley, M. Oke, N. Mizuno, J. Gumbart, E. Boura, A. N. Steere, O. Zak, P. Aisen, E. Tajkhorshid, R. W. Evans, A. R. Gorringe, A. B. Mason, A. C. Steven, and S. K. Buchanan, "Structural basis for iron piracy by pathogenic *Neisseria*," *Nature*, vol. 483, pp. 53–58, 2012.

[25] D. D. Shultis, M. D. Purdy, C. N. Banchs, and M. C. Wiener, "Outer membrane active transport: Structure of the BtuB:TonB complex," *Science*, vol. 312, pp. 1396–1399, 2006.

[26] A. Gudmundsdottir, P. E. Bell, M. D. Lundrigan, C. Bradbeer, and R. J. Kadner, "Point mutations in a conserved region (tonb box) of *Escherichia coli* outer membrane protein btub affect vitamin b12 transport," *J. Biochem.*, vol. 171, no. 12, pp. 6526–6533, 1989.

[27] C. Bradbeer, P. R. Reynolds, G. M. Bauler, and M. T. Fernandez, "A requirement for calcium in the transport of cobalamin across the outer membrane of *Escherichia coli*," *J. Biol. Chem.*, vol. 261, pp. 2520–2523, 1986.

[28] N. Cadieux, N. Barekzi, and C. Bradbeer, "Observations on the calcium dependence and reversibility of cobalamin transport across the outer membrane of *Escherichia coli*," *J. Biol. Chem.*, vol. 282, pp. 34 921–34 928, 2007.

[29] N. Kuerka, B. van Oosten, J. Pan, F. A. Heberle, T. A. Harroun, and J. Katsaras, "Molecular Structures of Fluid Phosphatidylethanolamine Bilayers Obtained from Simulation-to-Experiment Comparisons and Experimental Scattering Density Profiles," *J. Phys. Chem. B*, vol. 119, pp. 1947–1956, 2015.

[30] X. Zhuang, J. R. Makover, W. Im, and J. B. Klauda, "A systematic molecular dynamics simulation study of temperature dependent bilayer structural properties," *(BBA)–Biomembr.*, vol. 1838, pp. 2520–2529, 2014.

[31] A. Botan, F. Favela-Rosales, P. F. J. Fuchs, M. Javanainen, M. Kandu, W. Kulig, A. Lamberg, C. Loison, A. Lyubartsev, M. S. Miettinen, L. Monticelli, J. Mtt, O. H. S. Ollila, M. Retegan, T. Rg, H. Santuz, and J. Tynkkynen, "Toward atomistic resolution structure of phosphatidylcholine headgroup and glycerol backbone at different ambient conditions," *J. Phys. Chem. B*, vol. 119, pp. 15 075–15 088, 2015.

[32] R. M. Venable, F. L. H. Brown, and R. W. Pastor, "Mechanical properties of lipid bilayers from molecular dynamics simulation," *Chem. Phys. Lipids*, vol. 192, pp. 60–74, 2015.

[33] X. Zhuang, E. M. Dvila-Contreras, A. H. Beaven, W. Im, and J. B. Klauda, "An extensive simulation study of lipid bilayer properties with different head groups, acyl chain lengths, and chain saturations," *BBA–Biomembr.*, vol. 1858, pp. 3093–3104, 2016.

[34] K. Pluhackova, S. A. Kirsch, J. Han, L. Sun, Z. Jiang, T. Unruh, and R. A. Bckmann, "A Critical Comparison of Biomembrane Force Fields: Structure and Dynamics of Model DMPC, POPC, and POPE Bilayers," *J. Phys. Chem. B*, vol. 120, pp. 3888–3903, 2016.

[35] M. Doktorova, D. Harries, and G. Khelashvili, "Determination of bending rigidity and tilt modulus of lipid membranes from real-space fluctuation analysis of molecular dynamics simulations," *Phys. Chem. Chem. Phys.*, vol. 19, pp. 16 806–16 818, 2017.

[36] C. Kandt, W. L. Ash, and D. Peter Tieleman, "Setting up and running molecular dynamics simulations of membrane proteins," *Methods*, vol. 41, pp. 475–488, 2007.

[37] E. Lindahl and M. S. P. Sansom, "Membrane proteins: Molecular dynamics simulations," *Curr. Opin. Struct. Biol.*, vol. 18, pp. 425–431, 2008.

[38] M. Chavent, A. L. Duncan, and M. S. Sansom, "Molecular dynamics simulations of membrane proteins and their interactions: From nanoscale to mesoscale," *Curr. Opin. Struct. Biol.*, vol. 40, pp. 8–16, 2016.

[39] Y. Wang, T. Zhao, D. Wei, E. Strandberg, A. S. Ulrich, and J. P. Ulmschneider, "How reliable are molecular dynamics simulations of membrane active antimicrobial peptides?" *BBA–Biomembr.*, vol. 1838, pp. 2280–2288, 2014.

[40] C. T. Lee, J. Comer, C. Herndon, N. Leung, A. Pavlova, R. V. Swift, C. Tung, C. N. Rowley, R. E. Amaro, C. Chipot, Y. Wang, and J. C. Gumbart, "Simulation-Based Approaches for Determining Membrane Permeability of Small Compounds," *J. Chem. Inf. Model.*, vol. 56, pp. 721–733, 2016.

[41] A. P. Lyubartsev and A. L. Rabinovich, "Force field development for lipid membrane simulations," *BBA–Biomembr.*, vol. 1858, pp. 2483–2497, 2016.

[42] D. Poger, B. Caron, and A. E. Mark, "Validating lipid force fields against experimental data: Progress, challenges and perspectives," *BBA–Biomembr.*, vol. 1858, pp. 1556–1565, 2016.

[43] C. J. Dickson, B. D. Madej, Å. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould, and R. C. Walker, "Lipid14: The AMBER lipid force field," *J. Chem. Theory Comput.*, vol. 10, pp. 865–879, 2014.

[44] J. P. M. Jambeck and A. P. Lyubartsev, "An extension and further validation of an all-atomistic force field for biological membranes," *J. Chem. Theory Comput.*, vol. 8, pp. 2938–2948, 2012.

[45] J. B. Klauda, R. M. Venable, J. A. Freites, J. W. OConnor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, and R. W. Pastor, "Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types," *J. Phys. Chem. B*, vol. 114, pp. 7830–7843, 2010.

[46] A. Kukol, "Lipid models for united-atom molecular dynamics simulations of proteins," *J. Chem. Theory Comput.*, vol. 5, pp. 615–626, 2009.

[47] N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, and W. F. van Gunsteren, "Definition and testing of the GROMOS force-field versions 54A7 and 54B7," *Eur. Biophys. J.*, vol. 40, pp. 843–856, 2011.

[48] D. Poger and A. E. Mark, "Lipid bilayers: The effect of force field on ordering and dynamics," *J. Chem. Theory Comput.*, vol. 8, pp. 4807–4817, 2012.

[49] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. MacKerell Jr., "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields," *J. Comput. Chem.*, vol. 31, pp. 671–690, 2010.

[50] R. B. Best, X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, and A. D. MacKerell Jr., "Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ_1 and χ_2 dihedral angles," *J. Chem. Theory Comput.*, vol. 8, pp. 3257–3273, 2012.

[51] S. Jo, T. Kim, V. G. Iyer, and W. Im, "CHARMM-GUI: A web-based graphical user interface for CHARMM," *J. Comput. Chem.*, vol. 29, pp. 1859–1865, 2008.

[52] S. Jo, J. B. Lim, J. B. Klauda, and W. Im, "CHARMM-GUI Membrane builder for mixed bilayers and its application to yeast membranes," *Biophys. J.*, vol. 97, pp. 50–58, Jul. 2009.

[53] E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R. M. Venable, J. B. Klauda, and W. Im, "CHARMM-GUI Membrane Builder toward realistic biological membrane simulations," *J. Comput. Chem.*, vol. 35, pp. 1997–2004, 2014.

[54] J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks, A. D. MacKerell, J. B. Klauda, and W. Im, "CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the charmm36 additive force field," *J. Chem. Theory Comput.*, vol. 12, pp. 405–413, 2016.

[55] T. Romo and A. Grossfield, "Loos: An extensible platform for the structural analysis of simulations," *31st Annual International Conference of the IEEE EMBS*, pp. 2332–2335, 2009.

[56] T. D. Romo, N. Leioatts, and A. Grossfield, "Lightweight object oriented structure analysis: Tools for building tools to analyze molecular dynamics simulations," *J. Comp. Chem.*, vol. 35, pp. 2305–2318, 2014.

[57] Y. Wang, P. R. L. Markwick, C. A. F. de Oliveira, and J. A. McCammon, "Enhanced lipid diffusion and mixing in accelerated molecular dynamics," *J. Chem. Theory Comput.*, vol. 7, pp. 3199–3207, 2011.

[58] K. A. Feenstra, B. Hess, and H. J. C. Berendsen, "Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems," *J. Comput. Chem.*, vol. 20, pp. 786–798, 1999.

[59] C. W. Hopkins, S. Le Grand, R. C. Walker, and A. E. Roitberg, "Long-time-step molecular dynamics through hydrogen mass repartitioning," *J. Chem. Theory Comput.*, vol. 11, pp. 1864–1874, 2015.

[60] R. Francesco and S. Martin, "Thermodynamics and kinetics of largetimestep molecular dynamics," *J. Comput. Chem.*, vol. 33, pp. 475–483, 2012.

[61] J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, "Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes," *J. Comput. Phys.*, vol. 23, pp. 327–341, 1977.

[62] S. Miyamoto and P. A. Kollman, "Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models," *J. Comput. Chem.*, vol. 13, pp. 952–962, 1992.

[63] B. Loubet, W. Kopec, and H. Khandelia, "Accelerating All-Atom MD Simulations of Lipids Using a Modified Virtual-Sites Technique," *J. Chem. Theory Comput.*, vol. 10, pp. 5690–5695, 2014.

[64] P. Bjelkmar, P. Larsson, M. A. Cuendet, B. Hess, and E. Lindahl, "Implementation of the CHARMM force field in GROMACS: Analysis of protein stability effects from correction maps, virtual interaction sites, and water models," *J. Chem. Theory Comput.*, vol. 6, pp. 459–466, 2010.

[65] K. Olesen, N. Awasthi, D. S. Bruhn, W. Pezeshkian, and H. Khandelia, "Faster simulations with a 5 fs time step for lipids in the charmm force field," *J. Chem. Theory Comput.*, vol. 14, pp. 3342–3350, 2018.

[66] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, "Scalable molecular dynamics with NAMD," *J. Comp. Chem.*, vol. 26, pp. 1781–1802, 2005.

[67] D. Case, R. Betz, D. Cerutti, T. Cheatham, III, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao, and P. Kollman, *AMBER 2016*. University of California, San Francisco, 2016.

[68] M. J. Harvey, G. Giupponi, and G. D. Fabritiis, "ACEMD: Accelerating biomolecular dynamics in the microsecond time scale," *J. Chem. Theory Comput.*, vol. 5, pp. 1632–1639, 2009.

[69] A. A. Petruk, S. Varriale, M. R. Coscia, L. Mazzarella, A. Merlino, and U. Oreste, "The structure of the CD3 $\zeta\zeta$ transmembrane dimer in POPC and raft-like lipid bilayer: A molecular dynamics study," *(BBA) - Biomembr.*, vol. 1828, pp. 2637–2645, 2013.

[70] X. Liu, S. Ahn, A. W. Kahsai, K.-C. Meng, N. R. Latorraca, B. Pani, A. J. Venkatakrishnan, A. Masoudi, W. I. Weis, R. O. Dror, X. Chen, R. J. Lefkowitz, and B. K. Kobilka, "Mechanism of intracellular allosteric β_2AR antagonist revealed by X-ray crystal structure," *Nature*, vol. 548, pp. 480–484, 2017.

[71] D. Wacker, S. Wang, J. D. McCorvy, R. M. Betz, A. J. Venkatakrishnan, A. Levit, K. Lansu, Z. L. Schools, T. Che, D. E. Nichols, B. K. Shoichet, R. O. Dror, and B. L. Roth, "Crystal structure of an LSD-bound human serotonin receptor," *Cell*, vol. 168, 377–389.e12, 2017.

[72] M. Beeby, J. C. Gumbart, B. Roux, and G. J. Jensen, "Architecture and assembly of the Gram-positive cell wall," *Mol. Microbiol.*, vol. 88, pp. 664–672, 2013.

[73] L. T. Nguyen, J. C. Gumbart, M. Beeby, and G. J. Jensen, "Coarse-grained simulations of bacterial cell wall growth reveal that local coordination alone can be sufficient to maintain rod shape," *Proc. Natl. Acad. Sci. USA*, vol. 112, E3689–E3698, 2015.

[74] H. Hwang, N. Paracinic, J. M.Parks, J. H.Lakey, and J. C.Gumbart, "Distribution of mechanical stress in the escherichia coli cell envelope," *BJ*, vol. 1860, no. 12, pp. 2566–2575, 2018.

[75] C. Balusek and J. C. Gumbart, "Role of the native outer–membrane environment on the transporter BtuB," *Biophys. J.*, vol. 111, pp. 1409–1417, 2016.

[76] K. Lundquist and J. C. Gumbart, "Energetics of lateral gate opening in BamA," 2016, In preparation.

[77] K. Lundquist, C. Herndon, T. H. Harty, and J. C. Gumbart, "Accelerating the use of molecular modeling in the high school classroom with VMD lite," *Biochem. Mol. Biol. Educ.*, vol. 44, pp. 124–129, 2016.

[78] P. R. Reynolds, G. P. Mottur, and C. Bradbeer, "Transport of vitamin B12 in *Escherichia coli*," *J. Biol. Chem.*, vol. 255, no. 9, pp. 4313–4319, 1980.

[79] C. Bradbeer, "The proton motive force drives the outer membrane transport of cobalamin in *Escherichia coli*," *J. Bacteriol.*, vol. 175, pp. 3146–3150, 1993.

[80] D. M. Freed, S. M. Lukasik, A. Sikora, A. Mokdad, and D. S. Cafiso, "Monomeric TonB and the Ton box are required for the formation of a high-affinity transporter-TonB complex," *Biochemistry*, vol. 52, pp. 2638–2648, 2013.

[81] M. Kim, Q. Xu, D. Murray, and D. S. Cafiso, "Solutes alter the conformation of the ligand binding loops in outer membrane transporters," *BCH*, vol. 47, no. 2, pp. 670–679, 2008.

[82] Q. Xu, M. Kim, K. W. David Ho, P. Lachowicz, G. E. Fanucci, and D. S. Cafiso, "Membrane hydrocarbon thickness modulates the dynamics of a membrane transport protein," *BJ*, vol. 95, 2849–2858, 6 2008.

[83] D. P. Chimento, A. K Mohanty, R. J. Kadner, and M. C. Wiener, "Substrate-induced transmembrane signaling in the cobalamin transporter BtuB," *Nat. Struct. Biol.*, vol. 10, pp. 394–401, 2003.

[84] W. Humphrey, A. Dalke, and K. Schulten, "VMD – Visual Molecular Dynamics," *J. Mol. Graphics*, vol. 14, pp. 33–38, 1996.

[85] A. D. Ferguson, E. Hofmann, J. W. Coulton, K. Diederichs, and W. Welte, "Siderophore-mediated iron transport: Crystal structure of FhuA with bound lipopolysaccharide," *Science*, vol. 288, pp. 2215–2220, 1998.

[86] A. D. Ferguson, W. Welte, E. Hofmann, B. Lindner, O. Holst, J. W. Coulton, and K. Diederichs, "A conserved structural motif for lipopolysaccharide recognition by procaryotic and eucaryotic proteins," *Structure*, vol. 8, no. 6, pp. 585–592, 2000.

[87] A. D. Ferguson, J. W. Coulton, K. Diederichs, W. Welte, V. Braun, and H. P. Fiedler, "Crystal structure of the antibiotic albomycin in complex with the outer membrane transporter FhuA," *Protein Sci.*, vol. 9, no. 5, pp. 956–963, 2000.

[88] J. Stone, J. Phillips, P. Freddolino, D. Hardy, L. Trabuco, and K. Schulten, "Accelerating molecular modeling applications with graphics processors," *JCC*, vol. 28, pp. 2618–2640, 2007.

[89] T. A. Darden, D. M. York, and L. G. Pedersen, "Particle mesh Ewald: An N log N method for Ewald sums in large systems," *J. Chem. Phys.*, vol. 98, pp. 10 089–10 092, 1993.

[90] T. J. Piggot, A. Piñeiro, and S. Khalid, "Molecular dynamics simulations of phosphatidylcholine membranes: A comparative force field study," *J. Chem. Theory Comput.*, vol. 8, pp. 4593–4609, 2012.

[91] J. B. Klauda, B. R. Brooks, and R. W. Pastor, "Dynamical motions of lipids and a finite size effect in simulations of bilayers," *J. Chem. Phys.*, vol. 125, p. 144 710, 2006.

[92] J. Ulmshneider and M. Ulmschneider, "United atom lipid parameters for combination with the optimized potentials for liquid simulations all-atom force field," *J. Chem. Theory Comput.*, vol. 5, pp. 1803–1813, 2009.

[93] T. P. Straatsma and T. A. Soares, "Characterization of the outer membrane protein OprF of *Pseudomonas aeruginosa* in a lipopolysaccharide membrane by computer simulation," *Proteins: Struct., Func., Bioinf.*, vol. 74, pp. 475–488, 2009.

[94] A. H. Delcour, "Outer membrane permeability and antibiotic resistance," *Biochim. Biophys. Acta – Proteins and Proteomics*, vol. 1794, no. 5, 2009.

[95] L. A. Clifton, M. W. A. Skoda, E. L. Daulton, A. V. Hughes, A. P. LeBrun, J. H. Lakey, and S. A. Holt, "Asymmetric phospholipid: Lipopolysaccharide bilayers; a Gram-negative bacterial outer membrane mimic," *J. R. Soc. Interface*, vol. 10, no. 89, 2013.

[96] R. P. Dias, G. C. A da Hora, M. Ramstedt, and T. A. Soares, "Outer membrane remodeling: The structural dynamics and electrostatics of rough lipopolysaccharide chemotypes," *J. Chem. Theory Comput.*, vol. 10, no. 6, pp. 2488–2497, 2014.

[97] J. S. G. Slusky and J. R. L. Dunbrack, "Charge asymmetry in the proteins of the outer membrane," *Bioinformatics*, vol. 29, pp. 2122–2128, 2013.

[98] Martin B. Ulmschneider and Mark P. Sansom, "Amino acid distributions in integral membrane protein structures," *Biochim. Biophys. Acta*, vol. 1512, 1 2001.

[99] M. Schindler, M. Osborn, and D. Koppel, "Lateral diffusion of lipopolysaccharide in the outer membrane of *Salmonella typhimurium*," *Nature*, vol. 285, pp. 261–263, 1980.

[100] Shawkat, S. and Karima, R. and Tojo, T. and Tadakuma, H. and Saitoh, S. and Akashi-Takamura, S. and Miyake, K. and Funatsu, T. and Matsushima, K., "Visualization of the molecular dynamics of lipopolysaccharide on the plasma membrane of Murine macrophages by total internal reflection fluorescence microscopy," *J. Biol. Chem.*, vol. 283, 2008.

[101] Triantafilou, K. and Triantafilou, M. and Ladha, S. and Mackie, A. and Dedrick, R.L. and Fernandez, N. and Cherry, R., "Flourescence recovery after photbleaching reveals that LPS rapidly transfers from CD14 to hsp70 and hsp90 on the cell membrane," *J. Cell Sci.*, vol. 114, pp. 2535–2545, 2001.

[102] S. Snyder, D. Kim, and T. McIntosh, "Lipopolysaccharide bilayer structure: Effect of chemotype, core mutations, divalent cations, and temperature.," *Biochemistry*, vol. 38, pp. 10 758–67, 33 1999.

[103] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *JCP*, vol. 79, no. 2, pp. 926–935, 1983.

[104] C. Hong, D. P. Tieleman, and Y. Wang, "Microsecond molecular dynamics simulations of lipid mixing," *Langmuir*, vol. 30, no. 40, pp. 11 993–12 001, 2014.

[105] E. Yamashita, M. V. Zhalnina, S. D. Zakharov, O. Sharma, and W. A. Cramer, "Crystal structures of the OmpF porin: function in a colicin translocon," *EMBO J.*, vol. 27, no. 15, pp. 2171–2180, 2008.

[106] S. Pezeshki, C. Chimerel, A. N. Bessonov, M. Winterhalter, and U. Kleinekathofer, "Understanding ion conductance on a molecular level: an all-atom modeling of the bacterial porin OmpF," *Biophys. J.*, vol. 97, no. 7, pp. 1898–1906, 2009.

[107] P. S. Crozier, D. Henderson, R. L. Rowley, and D. D. Busath, "Model channel ion currents in NaCl-extended simple point charge water solution with applied-field molecular dynamics," *Biophys. J.*, vol. 81, no. 6, pp. 3077–3089, 2001.

[108] A. Aksimentiev and K. Schulten, "Imaging alpha-hemolysin with molecular dynamics: ionic conductance, osmotic permeability, and the electrostatic potential map," *Biophys. J.*, vol. 88, no. 6, pp. 3745–3761, 2005.

[109] W. Humphrey, A. Dalke, and K. Schulten, "VMD – Visual Molecular Dynamics," *J. Mol. Graphics*, vol. 14, pp. 33–38, 1996.

[110] A. R. Braun, E. G. Brandt, O. Edholm, J. F. Nagle, and J. N. Sachs, "Determination of electron density profiles and area from simulations of undulating membranes," *Biophys. J.*, vol. 100, no. 9, pp. 2112–2120, 2011.

[111] A. Kohlmeyer and J. Vermaas, *Akohlmey/topotools: Release 1.7*, Apr. 2017.

[112] H. Martinez-Seara and T. Róg, *Biomolecular Simulations. Methods in Molecular Biology, Methods and Protocols*. Humana Press, 2013, vol. 924, ch. Molecular Dynamics Simulations of Lipid Bilayers: Simple Recipe of How to Do It.

[113] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, "Mdanalysis: A toolkit for the analysis of molecular dynamics simulations," *J. Comput. Chem.*, vol. 32, no. 10, pp. 2319–2327, 2011.

[114] A. Grossfield and D. M. Zuckerman, "Quantifying uncertainty and sampling quality in biomolecular simulations," *ARCC*, vol. 5, pp. 23–48, 2009.

[115] H. I. Petrache, S. W. Dodd, and M. F. Brown, "Area per lipid and acyl length distributions in fluid phosphatidylcholines determined by 2H NMR spectroscopy," *BJ*, vol. 79, no. 6, pp. 3172–3192, 2000.

[116] N. Kučerka, S. Tristram-Nagle, and J. F. Nagle, "Closer look at structure of fully hydrated fluid phase DPPC bilayers," *BJ*, vol. 90, no. 11, pp. L83–L85, 2006.

[117] N. Kučerka, M. P. Nieh, and J. Katsaras, "Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature," *BBA*, vol. 1808, pp. 2761–2771, 11 2011.

[118] E. Egberts, S. Marrink, and H. J. C. Berendsen, "Molecular dynamics simulation of a phospholipid membrane," *EBPJ*, pp. 423–436, 22 1994.

[119] J. Wohlert and O. Edholm, "Dynamics in atomistic simulations of phospholipid membranes: Nuclear magnetic resonance relaxation rates and lateral diffusion," *JCP*, vol. 125, pp. 204 703–1–10, 2006.

[120] L. S. Vermeer, B. L. de Groot, V Réat, A. Milon, and J. Czaplicki, "Acyl chain order parameter profiles in phospholipid bilayers: Computation from molecular dynamics simulations and comparison with ^{2}H NMR experiments," *EBJ*, vol. 36, pp. 919–931, 2007.

[121] M. Camley B.A. Lerner, R. Pastor, and F. Brown, "Strong influence of periodic boundary conditions on lateral diffusion in lipid bilayer membranes.," *J. Chem. Phys.*, vol. 143, p. 243 113, 24 2015.

[122] G. Orädd, P. W. Westerman, and G. Lindblom, "Lateral diffusion coefficients of separate lipid species in a ternary raft-forming bilayer: A pfg-nmr multinuclear study," *BJ*, vol. 89, pp. 315–320, 2005.

[123] G. Lindblom and G. Orädd, "Lipid lateral diffusion and membrane heterogeneity," *Biochim. Biophys. Acta*, vol. 1788, pp. 234–244, 1 2009.

[124] W. Shinoda, N. Namiki, and S. Okazaki, "Molecular dynamics study of a lipid bilayer: Convergence, structure, and long-time dynamics," *JCP*, vol. 106, no. 13, pp. 5731–5743, 1997.

[125] M. Rappolt, A. Hickel, F. Bringezu, and K. A. Lohner, "Mechanism of the lamellar/inverse hexagonal phase transition examined by high resolution x-ray diffraction," *BJ*, vol. 84, no. 5, pp. 3111–3122, 2003.

[126] R. P. Rand and V. A. Parsegian, "Hydration forces between phospholipid bilayers," *BBARM*, vol. 988, no. 3, pp. 351–376, 1989.

[127] N. Kučerka, J. F. Nagle, J. N. Sachs, S. E. Feller, J. Pencer, A. Jackson, and J. Katsaras, "Lipid bilayer structure determined by the simultaneous analysis of neutron and x-ray scattering data," *BJ*, vol. 95, no. 5, 2356–2367, 2008.

[128] N. Kučerka, S. Tristram-Nagle, and J. F. Nagle, "Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains," *JMEB*, vol. 208, no. 3, pp. 193–202, 2005.

[129] W. Rawicz, K. C. Olbrich, T. McIntosh, D. Needham, and E. Evans, "Effect of chain length and unsaturation on elasticity of lipid bilayers," *BJ*, vol. 79, pp. 328–339, 2000.

[130] F. Heinemann, V. Betaneli, F. A. Thomas, and P. Schwille, "Quantifying lipid lateral diffusion by fluorescence correlation spectroscopy: A critical treatise," *LM*, vol. 28, no. 37, pp. 13 395–13 404, 2012.

[131] R. W. Benz, F. Castro-Román, D. J. Tobias, and S. H. White, "Experimental validation of molecular dynamics simulations of lipid bilayers: A new approach," *BJ*, vol. 88, no. 2, pp. 805–817, 2005.

[132] W. D. Bennett, C. K. Hong, Y. Wang, and D. P. Tieleman, "Antimicrobial peptide simulations and the influence of force field on t he free energy for pore formation in lipid bilayers," *J. Chem. Theory Comput.*, vol. 12, no. 9, pp. 4524–4533, 2016.

[133] E. G. Brandt, A. R. Braun, J. N. Sachs, J. F. Nagle, and O. Edholm, "Interpretation of fluctuation spectra in lipid bilayer simulations," *Biophys. J.*, vol. 100, no. 9, pp. 2104–2111, 2011.

[134] A. R. Braun, E. G. Brandt, O. Edholm, J. F. Nagle, and J. N. Sachs, "Determination of electron density profiles and area from simulations of undulating membranes," *Biophys. J.*, vol. 100, no. 9, pp. 2112–2120, 2011.

[135] D. Marsh, "Elastic curvature constants of lipid monolayers and bilayers," *Chem. Phys. Lett.*, vol. 144, no. 2, pp. 146–159, 2006.

[136] S. W. Cowan, R. M. Garavito, J. N. Jansonius, J. A. Jenkins, R. Karlsson, N. Konig, E. F. Pai, R. A. Pauptit, P. J. Rizkallah, and J. P. Rosenbusch, "The structure of OmpF porin in a tetragonal crystal form," *Structure*, vol. 3, no. 10, pp. 1041–1050, 1995.

[137] M. Watanabe, J. Rosenbusch, T. Schirmer, and M. Karplus, "Computer simulations of the OmpF porin from the outer membrane of Escherichia coli," *Biophys. J.*, vol. 72, no. 5, pp. 2094–2102, 1997.

[138] J. C. Gumbart, M. B. Ulmschneider, A. Hazel, S. H. White, and J. P. Ulmschneider, "Computed free energies of peptide insertion into bilayers are independent of computational method," *J. Mol. Biol.*, vol. 251, no. 3, pp. 345–356, 2018.

[139] H. Hwang, P. Yunker, and J. C. Gumbart, "Mechanics of bacterial inner and outer membranes," 2016, In preparation.

[140] G. Zhao, J. Perilla, E. L. Yufenyuy, X. Meng, B. Chen, J. Ning, J. Ahn, A. M. Gronenborn, K. Schulten, C. Aiken, and P. Zhang, "Mature HIV-1 capsid structure by cryoelectron microscopy and all-atom molecular dynamics," *Nature*, vol. 497, 643 EP, 2013.

[141] M. Sener, J. Strumpfer, J. Timney, A. Freiberg, C. N. Hunter, and K. Schulten, "Photosynthetic vesicle architecture and constraints on efficient energy harvesting," *Biophys. J.*, vol. 99, pp. 67–75, 2010.

[142] M. Feig, R. Harada, T. Mori, I. Yu, K. Takahashi, and Y. Sugita, "Complete atomistic model of a bacterial cytoplasm for integrating physics, biochemistry, and systems biology," *J. Mol. Graph. Model.*, vol. 58, pp. 1–9, 2015.

[143] A. Pavlova, H. Hwang, K. Lundquist, C. Balusek, and J. C. Gumbart, "Living on the edge: Simulations of bacterial outer-membrane proteins," *Biochim. Biophys. Acta*, vol. 1858, pp. 1753–1759, 2016.

[144] I. Botos, N. Majdalani, S. J. Mayclin, J. G. McCarthy, K. Lundquist, D. Wojtowicz, T. J. Barnard, J. C. Gumbart, and S. K. Buchanan, "Structural and functional characterization of the LPS transporter LptDE from Gram-negative pathogens," *Structure*, vol. 24, pp. 965–976, 2016.

[145] C. Chen, H. Huang, and C. Wu, "Protein bioinformatics databases and resources," *Methods Mol. Biol.*, vol. 1558, pp. 3–39, 2017.

[146] A. Nightingale, R. Antunes, E. Alpi, B. Bursteinas, L. Gonzales, W. Liu, J. Luo, G. Qi, E. Turner, and M. Martin, "The proteins api: Accessing key integrated protein and genome information," *Nucleic Acids Res.*, vol. 45, W539–W544, 2017.

[147] G. Orfanoudaki and A. Economou, "Proteome-wide subcellular topologies of *E. coli* polypeptides database (STEPdb)," *Mol. Cell Proteomics*, vol. 13, pp. 3674–3687, 12 2014.

[148] M. Wang, C. Herrmann, M. Simonovic, D. Szklarczyk, and C. von Mering, "Version 4.0 of paxdb: Protein abundance data, integrated across model organisms, tissues, and cell-lines.," *Proteomics*, vol. 15, pp. 3163–3168, 18 2015.

[149] L. Arike, K. Valgepea, L. Peil, R. Nahku, K. Adamberg, and R. Vilu, "Comparison and applications of label-free absolute proteome quantification methods on *E. coli*," *J. Proteomics*, vol. 75, pp. 5437–5448, 17 2012.

[150] J. R. Wiśniewski and D. Rakusb, "Quantitative analysis of the *Escherichia coli* proteome," *Data Brief*, vol. 1, pp. 7–11, 2014.

[151] J. R. Wiśniewski, M. Hein, J. Cox, and M. Mann, "A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards," *Mol. Cell Proteomics*, vol. 13, pp. 3497–3506, 12 2014.

[152] K. Krug, A. Carpy, G. Behrends, K. Matic, N. Soares, and B. Macek, "Deep coverage of the *Escherichia coli* proteome enables the assessment of false discovery rates in simple proteogenomic experiments," *Mol. Cell Proteomics*, vol. 12, pp. 3420–3430, 11 2013.

[153] F. Samsudin, M. Ortiz-Suarez, T. Piggot, P. Bond, and S. Khalid, "OmpA: A flexible clamp for bacterial cell wall attachment," *Structure*, vol. 24, pp. 2227–2235, 12 2016.

[154] J. Spector, S. Zakharov, Y. Lill, O. Sharma, W. Cramer, and K. Ritchie, "Mobility of BtuB and OmpF in the *Escherichia coli* outer-membrane: Implications for dynamic formation of a translocon complex," *Biophys. J.*, vol. 99, pp. 3880–3886, 2010.

[155] R. Koebnik, K. P. Locher, and P. Van Gelder, "Structure and function of bacterial outer membrane proteins: Barrels in a nutshell," *Mol. Microbiol.*, vol. 37, pp. 239–253, 2 2000.

[156] K. A. Sochacki, I. A. Shkel, M. T. Record, and J. C. Weisshaar, "Protein diffusion in the periplasm of *E. coli* under osmotic stress," *Biophys. J.*, vol. 100, pp. 22–31, 2011.

[157] M. Sotomayor and K. Schulten, "Molecular dynamics study of gating in the mechanosensitive channel of small conductance MscS," *Biophys. J.*, vol. 87, pp. 3050–3065, 5 2004.

VITA

EDUCATION

Georgia Institute of Technology 2012-2018

Ph.D. Physics — Dec. 2018

Thesis: Pushing the Cell Envelope:

Simulations at the Gram-Negative Cellular Interface

M.S. in Physics — Dec. 2013

Sam Houston State University 2008-2012

B.S. in Physics & Mathematics — May 2012

Magna Cum Laude & Academic Distinction

AWARDS

Georgia Institute of Technology

Leadership Education & Development

Graduate Student Excellence Award for Leadership & Service — 2018

Center for Teaching & Learning

Physics Graduate Teaching Assistant of the Year — 2017

Graduate Teaching Assistant Award Finalist (for campus-wide TA of the Year) — 2016

School of Physics

Graduate Teaching Assistant of the Year — 2017

Amelio Travel Award, $2,000 — 2017

Bonnie B. and Charles K. Rice Jr. Fellowship

for Excellence in research and the classroom — 2016

Biophysical Society Education Committee Travel Award, $750	2017
Oxford Nanopore Technologies Travel Award, 500€	2017
Department of Education, G.A.A.N.N. Fellowship	
Molecular Biophysics Training Program, $19,500	Aug 2015 - May 2016
Molecular Biophysics Training Program, $19,500	Aug 2014 - May 2015
Yellow Jacket Archery Club	
Coach	Aug 2016–Dec 2018
Tournament Coordinator	May 2014–May2016
Member	August 2013 – December 2017
USA Archery Collegiate Archery Program	
Collegiate 3D Nationals Men's Recurve Champion	2017
Collegiate All-Region Team (South)	2015, 2016
Academic All-American	2015, 2016
All-American	2015
US Collegiate Archery	
All-American	2014
Academic All-American	2014

PEER-REVIEWED PUBLICATIONS

C. Balusek, H. Hwang. *Role of the native outer-membrane environment on the transporter BtuB*. **C. Balusek** and J. C. Gumbart. Biophys. J. 111(7). 2016.

Lateral Opening and Exit Pore Formation Are Required for BamA Function. N. Noinaj, A. J. Kuszak, **C. Balusek**, J. C. Gumbart, and S. K. Buchanan. Structure. 22(7). 2014.

A mathematical model of biocontrol of invasive aquatic weeds. J. Alford, *C. Balusek*, K. Bowers, and C. Hartnett. Involve, J. of Math. 5(4). 2013.

A three-dimensional finite element model of near-field scanning microwave microscopy. **C.**

Balusek, B. Friedman, D. Luna, B. Oetiker, A. Babajanyan, and K. Lee. J. Appl. Phys. 112(8). 2012.

REVIEWS

Living on the edge: Simulations of bacterial outer-membrane proteins. A. Pavlova, H. Hwang, K. Lundquist, **C. Balusek**, and J. C. Gumbart. Biophys. Biochem. Acta. 1858. 2016